

Conversation analytic study of non-straightforward answers in political debates:

“First of all, I have to say, as a businessman, I get along with everybody”

Eeli Eilola  
Master's Thesis  
English Philology  
Faculty of Humanities  
University of Oulu  
Spring 2019

## Abstract

This thesis is a study of two different political debates which took place before the United States presidential election of 2016: a Republican debate that took place in South Carolina and a Democratic debate that took place in Nevada. The aim of this study was to discover how *well/well, first of all/first of all*-prefaced responses are used during the debates. The method used in this study was Conversation Analysis that has been used before when analyzing institutional settings as a debate is.

During the initial phase of the research it was quickly discovered that there were three main subsets of *well/well, first of all/first of all*-prefaced responses which were named as follows: the addition (giving additional information before answering a question), 'not answering the question' or NATQ, and 'commenting on a prior question' or CPQ. All the subsets represented different non-straightforward responses and they were analyzed individually. Three types of features were observed when looking at the responses of each subset: is the response straightforward, is the response pair-conforming (does the response have the expected form) and is the response sufficient. It was revealed that none of the subsets were systematically similar in their form throughout the debates. However, multiple convergent features were discovered.

Additions were either used as a circumbendibus back to the original topic, a diversion that enables avoiding the question, or an additional information that amplifies the following statement. When the debaters performed a NATQ they either first spoke topically and moved on to a different subject, answered a different question and moved to a different subject, or blatantly skipped the question in its entirety. The third subset CPQ was used when the debaters really wanted to comment on something before giving an elaborative answer. They usually gave a brief answer, commented on a prior question and moved back to the original question to elaborate their short answer. However, there was a case where this exact form did not occur, thus, the form is only the most common form that occurs.

In addition to the findings on the verbal aspect of *well/well, first of all/first of all*-prefaced answers, it was revealed that topic shifts are often accompanied by different sorts of gesticulation. The debaters often performed for example beat gestures, pointing gestures or shifted their gaze during the shifts. This study could not go deep into the details of why this happened, but it would be an interesting topic for future research.

## Tiivistelmä

Tässä gradussa tarkastellaan kahta poliittista debattia ennen vuoden 2016 Yhdysvaltain presidentinvaaleja. Ensimmäinen debatti käytiin demokraattiehdokkaiden välillä Nevadassa ja toinen republikaaniehdokkaiden välillä Etelä-Carolinassa. Tämän gradun tavoitteena oli saada selville, miten *well/well, first of all/first of all*-alkuisia vastauksia hyödynnettiin debattien aikana. Metodina gradussa on käytetty keskusteluanalyysia, jota on käytetty ennenkin institutionaalisten vuorovaikutustilanteiden, joihin debattitkin kuuluvat, tutkimiseen.

Tutkielman alkuvaiheessa aineistoa tutkiessa selvisi pian, että vastaustyyppit on helppo jakaa kolmeen seuraavanlaiseen alakategoriaan: lisäys (lisäinformaation antamista ennen varsinaiseen kysymykseen vastaamista), kysymykseen vastaamatta jättäminen ja aiemmin esitetyn kysymyksen kommentoiminen. Kaikki alakategoriat edustivat erilaisia epäsuoria vastaustapoja kysymyksiin, ja jokainen kategoria täten analysoitiin erikseen. Alakategorioita tarkastellessa kiinnitettiin huomiota kolmeen eri ominaisuuteen: onko vastaus epäsuora, onko vastaus oletetun vastaustyyppin mukainen ja onko vastaus riittävän laaja. Työn edetessä paljastui, ettei yksikään alakategoria ollut muodoltaan täysin yhteneväinen kaikissa esiintyneissä tapauksissa. Yhteisiä piirteitä kuitenkin löytyi siitä huolimatta.

Lisäys-vastauksia käytettiin joko ketunlenkkeinä alkuperäiseen aiheeseen, harhautuksena jonka avulla onnistuttiin välttämään kysymykseen vastaaminen, tai sellaisen lisäinformaation antamisena, joka tuki tulevaa vastausta kysymykseen. Kun osallistujat jättivät vastaamatta kysymykseen, he joko vastasivat puhumalla yleisesti ja siirtyivät toiseen aiheeseen, vastasivat aivan eri kysymykseen ja siirtyivät eri aiheeseen, tai jättivät häikäilemättä vastaamatta koko kysymykseen. Aiempaa kysymystä kommentoitiin silloin, kun osallistuja halusi kommentoida jotain aiempaa kysymystä ennen kuin vastaisivat kysymykseen. Yleensä osallistuja antoi lyhyen vastauksen alkuperäiseen kysymykseen, sitten kommentoivat toista kysymystä ja lopulta palasivat tarkentamaan lyhyttä vastaustaan alkuperäiseen kysymykseen. Kuitenkin aineistosta löytyi esimerkki, jossa juuri tämä vastauksen muoto ei toistunut, joten edellä mainittu muoto on se yleisin eikä suinkaan sääntö.

Verbaaliseen puoleen kuuluvien löydösten lisäksi paljastui, että puheenaiheen vaihdoksien aikana osallistujat hyödynsivät paljon visuospatiaalisia keinoja. Osallistujat esimerkiksi tekivät lyöntiliikkeitä, osoittivat ja siirsivät katsettaan näiden puheenaiheen vaihdoksien aikana. Tutkielmassa ei syvennytty ilmiöön sen enempää, mutta se voisi olla hyvä aihe tuleville tutkimuksille.

## Table of contents

1. Introduction.....	4
2. Theory and method .....	7
2.1. Conversation analysis as a theory .....	7
2.1.1. Basic concepts of CA.....	7
2.1.2. Intersubjectivity .....	10
2.1.3. Previous studies concerning political communication .....	13
2.1.4. Visuospatial modality .....	16
2.2. Conversation analysis as a method .....	20
3. Data .....	23
3.1. Data .....	23
3.2. Gathering data in an institutional setting.....	26
4. Analysis .....	30
4.1. 'Addition' .....	30
4.2. 'Not answering the question' .....	43
4.3. 'Commenting on a prior question' .....	53
5. Discussion and conclusion .....	63
References .....	68
Appendices .....	72

## 1. Introduction

Politics affects all of us whether we like it or not. Its undeniable ability to influence all of our lives without demanding all of our participation makes it a topic of major interest. Therefore, that is probably why researching such a powerful topic has been an interest of many other researchers before me within all areas of research.

Researchers practicing discourse analysis have been familiar with the topic of politics for decades. Analysis on political philosophy or research on political issues have been done before (Essed, Farquharson, Pillay & White, 2018; Braybrooke and Rice, 2006). In addition, guidebooks on how to analyze political discourse have been published before, as well (Chilton, 2004). Thus, generally the field of discourse analysis is familiar with the topic of politics. However, I would argue that there has not yet been detailed enough analyses on how the politicians act.

Conversation analysis, here on referred to as CA, is the area of research that provides the necessary tools for detailed analysis that will not interpret the results any further than what is observable. Using CA as the method provides detailed transcripts for determining patterns in politicians' behavior. Identifying how politicians utilize modalities in specific circumstances and detecting recurring patterns is something that could benefit not only conversation analysts, but CA can act as a complementary tool for other areas of researcher.

CA has been utilized widely when it comes to institutional talk in political settings. There are examples such as Poggi's (2013) research on multimodal communication in political speeches, Heritage and Clayman (2010) analyzing lengths of applause, Piirainen-Marsh (2005) analyzing television interviews and Atkinson (1984) discussing language and body language of politics. Debates have been analyzed by Hutchby & Allan (2006), however, in his study the focus was on rhetoric and what are the "practical and skilful means by which they (politicians) constructed alignments in the light of others' responses" (p. 159) used during a non-campaigning debate between representatives of opposing parties. However, in this study the focus is not on rhetoric but on responding in a live debate setting. Answering questions is the focus of debates and it is in the debaters' interest to show the audience their knowledge on topics and to persuade the listeners to their side. The optimal way to do it is to answer the judges' questions or avoid bad ones and to be

able to answer questions in a manner that maximizes the number of new supporters gained and/or minimizes the number of supporters lost. Thus, the politicians must make a fast assessment of the question asked if it is something that should be answered straightforwardly or should one try to dodge it.

Furthermore, the participants in the debates are running for the nomination of their own parties, and therefore it is reasonable to argue that this study breaks new ground as it is something that has not yet been analyzed, in comparison to interviews and debates against candidates of different parties have been. This type of CA research could benefit all the participants both future debate judges and debaters. The study could reveal for example when politicians are trying to move away from the subject using a specific line of dialogue or other means of conversation. In addition, it could be beneficial from the politicians' perspective to know if there are techniques that when answering a challenging question are more prone to initiate successful diversion.

What makes a live debate a noteworthy subject for studying is that a debate is a situation where the participants are stationary. They do not have the option of leaving the space if things turn out too heated. Live debates let the audience see journalists asking tough questions from politicians and the politicians cannot physically escape the situation. Therefore, their answers and different actions around those answers reveal patterns in speech related to the live debate situation. Revealing patterns from politicians' responses would be a great illustration of how CA can be utilized in studying political language. In this study I will examine two live recordings. First one is the recording of a South Carolina Republican Debate recorded by CBS in early 2016, where tough questions are presented to the debaters Ben Carson, Marco Rubio, Donald Trump, Ted Cruz, Jeb Bush and John Kasich by Judge (either John Dickerson, Major Garrett or Kimberley Strassel). The second recording is the recording of a Nevada Democratic Debate recorded by CNN in late 2015, where the questions were presented to the debaters Lincoln Chafee, Hillary Clinton, Martin O'Malley, Bernie Sanders and Jim Webb by Judge (either Anderson Cooper, Dana Bash, Don Lemon or Juan Carlos López).

After going through the debates, one very frequent way of answering questions was to preface the response with either *well*, *well first of all* or *first of all*, hence *well/first of all*, to a variety of different questions. The phrases seemed to have no tonal difference in their context and therefore they are mashed together as the phrase of interest in this study. The phrases were accompanied with sort of *non-answers*, referring to anything but a straightforward answer, that exemplified politicians' abilities of speaking topically or avoiding a question completely. Sometimes the answers were

followed by either a follow-up question or a pursuit that portrayed how the judges were not always content with the answer they were given.

First, I will discuss CA as a theory and method for conducting research and then I will describe the data in more detail in section 3. Section 3 will be followed by the analysis of the data in section 4 which will lead to Discussion and Conclusion in section 5. Discussion and Conclusion will be followed by References and Appendices respectively.

## 2. Theory and method

Conversation analysts seek to understand human interaction and what the means utilized in achieving goals in communication are. Conversation analysts pursue these goals by doing descriptive analysis on verbal and non-verbal communication between participants. The theory has evolved as more and more of analysis on communication has been conducted and at the same time, due to researchers not living in complete isolation of the outside world, CA has drawn influence from other research, as well. However, it is not a question of “which came first: the theory or the method?” but rather how have they complemented each other. Therefore, it does not really matter where you start as long as the link between the two topics becomes clear. This section starts with theory as in my opinion it feels more intuitive and gives sensible background for the topic of CA.

### 2.1. Conversation analysis as a theory

There is a wide variety of concepts and terms in CA. The theory has evolved further and further as more research has been done by utilizing the theory. Therefore, I have decided to discuss only couple of the key concepts of CA that I think are relevant to this study starting with the concept of ‘action’.

#### 2.1.1. Basic concepts of CA

According to Sidnell (2010) conversation is built of *turns* that perform *actions*. Turns are what we say during the time we have, and an action is a goal of communication. Language itself does not accomplish anything, but it is the tool utilized in accomplishing actions. While there are many ways or *practices* that can accomplish the same action, actions are inseparable from practices in the way that every practice aims to accomplish an action. For example, there are multiple ways one might try to greet another person they have just met, but some of the possible practices might be more effective than others. Saying “hello” to someone and looking them into their eyes might be more



effective than yelling “hello” next to them and rolling one’s eyes. Both have the same verbal component “hello”, but the verbal component is accompanied by different kind of tone and eye movement. In addition, context matters as in western societies people do not tend to yell to each other when meeting for the first time. Yelling “hello” could mean something else such as trying to get someone’s attention. In a nutshell, turn is the phase where one says “hello”, greeting is the action, and the way one performs the action is the practice.

Sidnell (2010) says that in CA researchers should be able to identify these types of practices from data. Initially one’s intuition helps when trying to find appropriate data, but then evidence must be gathered. There are three types of evidence for researchers to gather. First, researchers should try to identify recurring examples of a construction that tries to achieve the focal action. Second, researchers should look at “what comes next in the sequence” and see how the participants react to the utterance in focus. Finally, the researcher should “look at the relation of the focal turn and the previous one” and see if the turn itself was “designed as a response.” (pp. 60-63) These guidelines are kept in mind along this study. Gathering evidence of all the types presented is preferable as identifying or naming actions should not be done haphazardly. After actions are identified they can be analyzed as sequences if necessary. The most basic sequences consist of two actions and it is called an *adjacency pair*.

Schegloff (1968) had expressed his concern with the term adjacency pair. The term seemed to imply that the actions are an inseparable pair of conversation. However, sometimes actions of a such pair can occur without the other. (Schegloff, 1968, as cited in Sidnell, 2010, p. 63) For the case of clarity, a single type of adjacency pair is discussed which is “question and answer.” A question is not always followed by an answer which leads to the conclusion that adjacency pairs are not a rule but what is expected by participants. When a question is uttered, the participants will assess whether the following action can be identified as an answer. An answer is the expected follow-up, but it is not always given that the following action is an answer. The organization of adjacency pairs is not a probability. As Sidnell (2010) says: “rather, the organization described is a norm to which participants hold one another accountable” (p. 64). The normative organization is born from the expectation of participants’; the expected becomes the norm in conversation. Therefore, a noteworthy subject of studying related to adjacency pairs is absence of an answer or reply. Diverting the expectation of the other participant by not giving the normative answer is at least as interesting as studying normative responses. Often the reason for absence of answer is not knowing, but

because there is such a wide variety of different reasons for absence of answer, CA does not try to define a “proper” way to answer by looking at probabilities. In addition to “answer and question” there are numerous types of different adjacency pairs with numerous types forms. All the adjacency pair types and their forms of adjacency pairs will not be listed here, but they will be discussed when they are apparent in the data gathered.

Another type of adjacency pair apparent in this study is the imperative form of a request which is paired with a response. The requests that matter in this study are imperatively formatted directives, where an actor directs another actor to do something by using an imperative form of a verb. Imperative formatted directives have been studied by Etelämäki & Couper-Kuhlen (2017) who were concerned with how in adult conversation actors resisted imperative directives and how the initiator of the directive could act in the situation, which was by dropping the matter or to re-issue the directive (pp. 215-216). However, in this study the participants either comply to directives and deliver an expected response or stay out of the topic completely. The most effective way to not raise further directives is either responding or responding in a manner that is masked well as a response. The latter of the two is closely related to the concept of *pursuit* which will be discussed in Section 2.1.3.

When utilizing CA researchers should not try foreseeing human actions, because CA is not deterministic in nature. What CA is for is to describe what kinds of actions humans take during conversation (Sidnell, 2010, pp. 65-66). Foreseeing actions in sequences of two turns is sometimes possible, but because most of the time sequences tend to be longer and more complex, foreseeing becomes impossible quite soon on the line. Parts of sequences called expansions make the schematics of interaction much more difficult to foresee or chart through, because they can occur “before the occurrence of its first part, after the occurrence of its first part but before the occurrence of its second, or after its second part.” The three expansions are called pre-expansion, insert expansion and post expansion. (Sidnell, 2010, p. 95) Pre-expansions typically “check on a condition for the successful accomplishment of the base first-pair part” (Sidnell, 2010, p. 95) and pre-expansions sometimes have the ability to skip the first-pair part completely, because the other participant can anticipate what the first-pair part, here on FPP, will be about. How this works can be seen from an example from Sidnell (2010) where Mike inquires if Vic has planned to do anything with a large tank and Vic realizes right away that Mike is planning to ask if the tank is for sale or for free:

- (1) US. 24  
 01 Mike: —→ Wuhddiyuh doing wh dat big bow-puh-tank.  
 02 Nothing?  
 03 (0.5)  
 04 Vic: ((cough))  
 05 Uh-h-h  
 06 (1.0)  
 07 Vic: —→ I'm not intuh selling or giving it. That's it.  
 08 Mike: Okay  
 09 (1.0)  
 10 Mike: Dat wz simple. Khhhh huh-huh-heh=  
 11 Vic: =Yeh. (Sidnell, 2010, p. 96)

We are able conclude that pre-expansions, in this case pre-request, have their own rules they follow and create which further on complicates the nature of conversation. Same applies to other expansions, too. The aspect of expansions will be discussed later on in the study when it is relevant in the analysis section.

### 2.1.2. Intersubjectivity

An important dimension of CA related to this study's topic concerns the relations between participants and the discussion at hand. During conversation the participants' realities intersect, preferably, and they share an idea of what the conversation is about and acknowledge the fact that the other participant is trying to convey something to them. Sidnell (2010) defines this phenomenon of "joint or shared understanding between persons" as *intersubjectivity* (p. 12). The term's roots are in philosophy. Ballard (1971) discussed German philosopher Husserl's ideas of intersubjective world in *Philosophy at the Crossroads* and according to Husserl there are two layers to the world. The more personal layer is the one where we as subjects see the world from our point of view. That is the way we view the world most of the time, because we are usually the only human actors in an environment and we have no need for another point of view than the "real" self. The second layer consists of all that can be perceived from other points of views at given time. We create the "other-like-me" for the situations where it is necessary. Ballard portrays the idea as me having a specific point of reference "here" and for the other the point is "there". However, I can constitute the noematic "other-like-me" which enables me to attribute him a "here" for which I am "there".

Ballard (1971) argues that creating the “other-like-me” is inevitable for human beings of high enough level of mental capacities. Therefore, a person with high enough level of cognitive abilities has to go through the mental process of rationalizing the other person’s position, which creates the intersubjectivity in any communicative situation. Naturally, there are some living but communicating forms of life who do not possess the high enough level of cognitive abilities. Very young babies and animals – although determining whether animal intersubjectivity exists or not has proven to be a more complicated issue than initially it would have seemed like (Racine, 2016) – do not seem to possess the form of intersubjectivity that we think adult human beings possess. Thus, it is relevant to study the lack of it in conversation where all the participants should possess the ability but choose not to wield it such as in political debates.

In addition to philosophy and CA, intersubjectivity is a widely discussed topic in other fields of research, too: psychology and cognitive science. A major problem with intersubjectivity as an idea is that it conflicts to some extent with popular cognitive theories “theory of mind,” “theory theory” and “simulation theory.” (Zlatev, 2008, pp. 2-4) The definition and the idea of the term intersubjectivity in CA originates from psychology, and therefore theory of mind has to be challenged. Theory of mind proposing that there is a primary separation between self and others and that the only way to bridge it would be to create a simulation or a theory of the other’s mind is perplexing. Zlatev (2008) critiqued the theory by saying that from the theory of mind “point of a departure, it is unsurprising that there appears to be not only a divide, but a veritable gulf between self and others, one that is so wide that it is doubtful whether it could ever truly be bridged” (p. 2). If there was such an enormous gap between actors, how would children ever be able to gain any knowledge of a shared language. There has to be something else between the actors than the gap suggested, because otherwise acquiring competence in any language would be insuperable for infants.

What comes into existence between participants in conversation is the *intersubjective field*. Peräkylä (2013) describes the term as “an emergent field of shared understandings regarding each other’s actions and the worlds of momentary experience that these actions embody” (p. 552). The field contains gaps, but most of all the participants’ fields come together at some points. Crossing the intersubjective field in conversation creates the shared understanding between the participants involved in the communicative act and initiates the appearance of adjacency pairs; without the shared understanding of what the participants’ actions embody, no reactions occur.

According to Robinson (2014) in the field of CA the common understanding of how intersubjectivity is managed was by reference to the next-turn talk. However, they challenge that view as oversimplification, because CA research has produced information that suggests “that participants do not rely exclusively on next-turn talk in order to manage intersubjectivity” (p. 110, 122) when assessing whether the other participant understood the message or not. According to Robinson (2014) participants may sometimes have to progress to the next turn without getting nothing more than silence as response to the initial turn on a transition-relevant place, which does not confirm that the participants share an understanding, but it is what the omission suggests. Turn constructional unit, hence TCU, of repairing do not always take place on transition-relevance places. According to Robinson (2014) “the rules of turn taking organize conversation on a TCU-by-TCU basis, with TCUs being vehicles for implementing actions” (p. 112), and initiating repair is not a mandatory action, as it is something that is often expected in particular settings, but not always required. Thus, Robinson (2014) claims that if recipients do not use the opportunity to initiate repair, the setting suggests that intersubjectivity had not been lost (p. 112).

For intersubjectivity to occur, the participants have to register the preceding utterances during the conversation. In conversation intersubjectivity manifests its existence most clearly when problematic (ambiguous, incoherent or some other way unclear) lines are uttered during a conversation. When one of the participants does not understand an utterance, intersubjectivity is often lost and participants do not cross each other on the intersubjective fields. Kitzinger (2013) gives an example where a mother asks her child what biscuits she bought, and the daughter misunderstood the question. She thought that her mother was asking what types of cookies, when in reality her mother was asking what make they were (pp. 247-248). During the conversation both participants thought they knew exactly what they were discussing about until the mother realized her daughter did not understand what she meant, and she had to clarify her question. Therefore, it is true that participants may at some point during the conversation think that they understand each other, when in reality intersubjectivity was lost already somewhere along the way. So, it would not be unreasonable to assume that it is the speaker’s burden to exercise as precise language as possible during conversation in order for intersubjectivity to occur during the conversation. However, that is not always the case.

Irony in conversation is a great example of how maintaining intersubjectivity is the recipient’s responsibility, as well. Sidnell (2010) discusses irony and presents that, because when speaking

ironically the words uttered signify the opposite, using irony is a sort of “a test for the maintenance of intersubjectivity” (p. 70). The recipient has to somehow verify to the speaker that they as a recipient has understood their words and what they actually meant. The recipient may laugh as a sign of understanding the humor or they can fabricate an ironic response of their own which preserves the ironic conversation. Thus, intersubjectivity is not solely dependent on participants’ ability to articulate themselves as clearly as possible, but also on how well the participants can understand each other as recipients. Therefore, conversation analysts should only depict what they can see in the data instead of trying to read the participants’ minds.

### 2.1.3. Previous studies concerning political communication

The topic of studying politicians’ answers in interviews has been studied by discourse analysts before. Studies where politicians’ responses to an accusation of inconsistency have been analyzed and evaluated before, but in those studies the focus was on describing responses’ forms, response practices and responses’ function. The studies were purely discourse oriented as their goal was to concentrate on the transmission of information. (Jucker, 1986; Emmertsen, 2007) According to Andone (2013), there are some other studies that have taken into account that interviews are to some extent argumentative situations and therefore the participants’ focus is on convincing the audience instead of convincing each other. However, Andone (2013) adds that “although these studies show a better understanding of how arguers typically behave in a political interview, they lack a systematic theoretically-situated analysis that could provide an insight into the argumentative function of a politician’s moves” (p. 12). Andone’s (2013) research tried to observe and verify functions for the politicians’ moves. It was revealed that “a suitable adjustment enabling the politician to continue the discussion is to reformulate one of the standpoints at issue” (p. 127). This was a way to maneuver around the accusation and trying to increase odds of “winning” the argument. Politicians used three different patterns to accomplish this. First one was to reformulate the standpoint by claiming that there are more conditions to fulfill before the politician supports the standpoint. Second one was to portray the interviewer’s standpoint as a false interpretation. Third pattern was to reformulate the original standpoint and make it seem like that it concerned with something different than it originally did. (Andone, 2013, pp. 127-128) Researching political

discourse is not only limited to discourse analysis, as it has been researched in CA, too. While discourse analysts focus on the responses' effectiveness and the reasons for them, conversation analysts are interested in the formation of actions.

First, we will discuss how conversation analysts have examined communication from the interviewers' side. It is important to know what actions interviewers, hence IR, take after they realize that the interviewee, hence IE, was avoiding a question. The viewpoint is necessary to take in to account as it works as a great tool for identifying when the IEs are not responding properly.

IRs' cannot force IEs to answer their questions in a straightforward manner, but they can utilize a variety of different practices, the most important of them being called *pursuing*, to perform the action of getting a satisfying answer. According to Romaniuk (2013), politicians' tendency to "resist, subvert, or sidestep the constraints of hostile questioning" (p. 145) has been a topic of many discussions and debates, but whether they do it purposely or not is not the core of the matter; IRs should feel incited to reveal the true motivations regardless of IEs' responses. It is the IRs' duty to get the answers out from the politicians on behalf of the audience. (pp. 145-147) Therefore, if a politician fails to answer a question, journalists might start pursuing the answer by making a follow-up. Pursuing is defined as asking "questions that topicalize an IE's (c)overt refusal to answer the IR's prior question and make that the focus of the IE's next turn" (Greatbatch, 1986, as cited in Romaniuk, 2013, p. 147). Pursuing is a powerful tool that can be used by IRs when they feel that an IE's unsatisfactory answer should be exposed, and a new answer should be put on a pedestal.

Following Sidnell's (2010) guidelines of gathering data, one should pay attention to what follows the turn of interest. Often the judges either asked further questions or repeated their question and pursued for the answer if the question was not answered. Romaniuk (2013) lists four ways of initiating pursuit follow-ups (all the examples in this paragraph are from their research). First way is to explicitly reference the initial question as Wolf Blitzer did while interviewing Hillary Clinton in 2007: "Let me just be precise, because the question was: Do you regret not reading the national intelligence estimate?" Here the IR shifts the focus back to the initial question and at the same time succeeds in revealing the IE's attempt to avert the conversation. The reference is explicit as the IR refers to it specifically by saying: "..., because the question was:" (Romaniuk, 2013, p. 149) Second way is the verbatim repeat where the IR repeats the initial question word for word. In the example the IR asks if the IE "threatened to overrule Mister Howard" which the IR repeats after the IE's insufficient response to the same exact wording of the question. (p.8) Third way is to run with the

indexically linked pursuit which is not as explicit, but nevertheless succeeds in for example assuring the standpoint of an IE. In an example of this case the initial question was “Do you favor the teaching of creationism in public schools?” which was followed up by the IR with “Is that a yes?” The IR is making sure that the answer was interpreted correctly and at the same time audience has to remember what the question was exactly. Thus, the pursuit is indexically linked (p. 151). The fourth way is to prepare an implicitly formulated pursuit, where the initial question is neither referred to nor is there any indexical link between the initial question and the pursuit. In the example the IR asked first if the IE was open for the idea of raising payroll task. After the IE’s response to that question, the IR asks: “What about increasing the retirement age from sixty-five?” However, the IE keeps on discussing the first question instead of answering the new one. When the IR realizes this, the IR asks again: “And how do you feel about raising the retirement age?” to which the IE finally answers. (pp. 152-153) Pursuits show that questions’ only purpose is not to get responses, but to get specific answers, too. Using pursuits are a favorable way of interviewing politicians, because it is audiences’ desire to know exactly what politicians are up to.

Schegloff and Lerner (2009) have studied *wh*-questions and *well*-prefaced answers to them. The particular line pursued in their research addressed to the occurrence of *well* in turn-initial position in the second-pair part positions in these sequences. Earlier studies had regularly come to the understanding that *well* was an indicator of dispreferred response. However, Schegloff and Lerner’s (2009) observations proved otherwise as according to them that conclusion does not cover all the instances of *well*-prefaced answers. They claim that a *well*-prefaced responses to the *wh*-questions suggests that the answer to follow is to some respect non-straightforward, but it is more of a call for attention to the response to follow than an indicator of non-straightforwardness (p. 101). Schegloff and Lerner (2009) also discuss what would be a type-conforming response to *wh*-questions and the problematic detail is that an answer to such question as *where* could be a reference to a location but depending on the context the answer may not be type-conforming, because circumstances were not taken into an account. Therefore, in this study the term pair-conforming is used as it does not take a stance on whether the answer was sufficient or not but depends only on if the response is adjacency pair-conforming or not.

According to Schegloff and Lerner (2009), all utterance format questions have the ability to initiate a sequence that tends to have clear sequence-conforming (pair-conforming in this study) second-pair part, here on SPP. From this point of view, it was easy for them to address if starting with a *well*-



prefacing is exclusive for answering. What they discovered was that making a type-conforming answer to a *wh*-question did not automatically mean that the answer was sufficient. The answer might be adjacency pair-conforming, but in order to it being a good enough answer, the local context should be taken into an account (p. 21-22). From the perspective of analyzing political debates this conclusion is remarkable. Answers that are pair-conforming but insufficient are common in debates as politicians try to get around questions. What shall be taken into account in this study is whether the answers are pair-conforming or not, because if they are not, it is a good indicator that some sort of question averting is taking place. Acknowledging that a pre-second expansion before the SPP, such as the preface *well*, does not signify anything in the context of whether an answer to follow is sufficient or not.

#### 2.1.4. Visuospatial modality

We might instinctively think that humans are far above other species when it comes to our language. We are able to refer to past events and plan for the future. Cohen, Comrie and Lefebvre (2013) suggest that because early primates evolved in arboreal environments, their limbs adapted to the environment around them, thus becoming longer. After humans branched off the other species, they started utilizing the bipedal posture. Now their hands were free to act in collaboration with their ability to move freely in the four-dimensional space-time. Using hands as the medium of communication, usually mimicking space-time events, was the easiest way of conveying messages to an audience. (pp. 172-173) Vocal communication has been said to emerge from these circumstances.

Only a couple of species have the ability of vocal learning, of which none is a close relative of the human species. After observing chimpanzees, researchers have concluded that chimpanzees have little control over their vocal modality. Naturally, there are some deviants to the rule such as when a chimpanzee is describing level of aggression facing them or when a chimpanzee is describing the predators facing their group (Slocombe, Kaller, Call & Zuberbühler, 2010; Ouattara, Lemasson & Zuberbühler, 2009, as cited in Cohen, Comrie & Lefebvre, 2013). However, when observing visuospatial means, mainly gestures, neurologists have said that chimpanzees can learn an approximation to a low-level human language through a mirroring mechanism. A mirroring neuron

activates in a chimpanzee's brain when another chimpanzee or human is interacting with an object. The neuron picks up not only visual cues but sounds as well, such as sound of nut cracking. Then the chimpanzee may learn the causality of the action and mirror the action themselves. Interestingly, while brain scans have proven that chimpanzees' neurons react only to some other chimpanzee or human interacting with an object, scans have also shown the major difference between humans and chimpanzees: human neuron activates even when the other individual is acting in the absence of an object. The difference between us and them is us mirroring intransitive actions. Thus, researchers have suggested that this ability has been the first step of gestures gradually evolving to a complete human language. (Cohen, 2009, pp. 174-176) The connection between gestures and human language is undisputable. Thus, Stivers and Sidnell's (2005) claim of visuospatial means having communicative import in interaction, in addition to having purpose in language production, is validated by research in evolutionary psychology. For that reason, gestures will be taken into an account in this study, too.

Even though all we can hear of each other's thoughts are inevitably chains of vocal utterances, communication is not solely dependent on the vocal modality. In face-to-face conversation there is a whole other world of communicative means that contributes to conversation: the visuospatial. Visuospatial modality consists of manual gestures, facial expressions and body posture. Together with the vocal modality, – that is prosody, lexis and grammar – they form *multimodal* interaction. Mondada (2016) argues that interaction can never be solely dependent on the vocal modality and no aspect of language is free of multimodality, because there is always embodiment involved. Most of the time (excluding people using dictation machines) people have to use their respiratory system in order to produce speech, which creates body movement and breathing. In addition, producing speech causes movement of one's face and mouth which creates articulatory movement. Therefore, studying only words spoken can never describe language production as completely as necessary. This change of perspective affects the way conversation analysts see communication (Mondada, 2016, pp. 340-341).

According to Mondada (2014a), seeing interaction as fundamentally multimodal enables to either focus on one resource in relation to everything else or the interplay of multiple resources. To be more precise Mondada (2014a) gave three arguments for the usefulness of the perspective as:

- (a) it allows one to treat linguistic and embodied resources in principle in the same way, without prioritizing a priori one type of resource over other ones;

(b) to identify not only conventionalized sets of resources, as grammar and some types of gesture, but also situatedly occasioned and assembled resources;

(c) to study how resources are combined together in various configurations, depending on the activity, its ecology and its material and cultural constraints (p. 139)

Considering these three notions, it is reasonable to claim that prioritization of what to study and what to discuss should be a product of a long process of observing data. Especially, because the initially similar looking gestures can be realized in multitude of different ways with minor changes in detail or the same action can be achieved by combining completely different resources. In Mondada's (2014a) research it was shown that pointing gestures, where the participant was using pen as an artefact, proved their malleability. The gesture was utilized as a resource for either turn construction, turn-taking management, sequential organization, or reference in achieving multiple different actions. There were various ways of holding the pen and what activities to perform with it. By combining these options, participants performed various actions with the same resources. In the same paper it was exemplified that a specific formatting of an action may be selected depending on context of an action. In the example a surgeon was dissecting a patient and giving orders to their assistant. Depending on the context, the surgeon utilized different resources to perform the same action. They could either give a verbal command or point to a location in order to get the participant to for example to change the positioning of their pliers (Mondada, 2014a, pp. 142-151).

Mondada (2016) claims that conversation analysts see interaction as multimodal which refers to "the various resources mobilized by participants for organizing their action" (p.338). The plurality of modalities suggests that the modalities are inherently intertwined, and therefore, none of them are hierarchically above one another. The focus of conversation analysts is not on the modalities itself, but on what is the sequential organization of different means and how participants format turns, create actions or create sequences of actions within social interaction (Mondana, 2016, pp. 337-339).

All the forms of visuospatial communication have been studied and it has been observed that they contribute to conversations as their own or in a particular sequential context. Stivers and Sidnell (2005) discuss the function of visuospatial modality and they claim that there are two types of use for the visuospatial modality. Through evidence they claim that certain types of means may be connected with language production and "strong evidence exists to suggest that gestures have basic

communicative import for interaction” (p. 4). The claim makes sense from an evolutionary perspective as the current view on language development suggests that gestures preceded vocal communication.

Whether participants’ visuospatial means are intentional or not cannot be determined from video data itself, because one cannot interpret participants’ thoughts. However, by looking at large amount of data researchers can make observations about the circumstances surrounding the means and therefore make assumptions of when the means in question are used. The principle works the same in both observing vocal and visuospatial means. For example, with body posture people can express their level of participation to a conversation by standing in an F-formation, a formation where people form a polygon and their bodies are turned towards the center, or “when conversing in a torqued body posture (i.e., one part of the body oriented in one direction and another in an alternative direction)” a speaker suggests that the activity is of shorter term or subordinate relative to the other activity at hand (Kendon, 1990, as cited in Stivers and Sidnell, 2005; Goffman, 1963, as cited in Stivers and Sidnell, 2005). In addition, visuospatial modality is not utilized only when actors are joined in a multi-participant conversation. Sometimes visuospatial abnormalities may suggest of some underlying medical disorders. Lausberg (2014) gathered studies concerning mental disorders and their symptoms, and how they affect the visuospatial modality in individuals. One of the symptoms of depression is a general reduction in body movement or an increase in body movement. Lausberg’s conclusion, which they collected from the studies referenced, was that due to the symptoms of depression, depressed individuals may have an unusually slumped posture when compared to people without depression. In addition, depressed individuals expressed far less vertical/sagittal movement in relation to horizontal/lateral in total when compared to the healthy control groups. Thus, the goal is to determine the circumstances not the thoughts of participants. Circumstances may sometimes be conscious choices, as the example of joining a conversation, or they can be unintentional circumstances as the symptoms of depression.

Using hands in communication has been studied considerably for example palm position, movement of the hand or hand’s position in the gesture space. Even something as specific as a recurring gesture “palm up open hand gesture” and which gestures it has been combined with, has been in focus of multiple studies. The “palm up open hand” position has been seen commonly used in conversations for example when the participant interrupting, listing or even when the participant is describing abstract on-going processes. (Kendon, 2004; Ladewig, 2014) Sometimes hand gestures may even

transform a particular utterance to convey a different meaning. Kendon's research in Southern Italy concluded that gestures such as the "Ring" and the "Finger Punch" can go far beyond their usual meaning depending on the context they are (Kendon, 1995).

According to Goodwin (1981) gaze often accomplishes the job of selecting a recipient in multiparty interaction or it can be utilized in showing other participants that one is focusing in the communicative act. (p. 30) A more recent study by Jarmolowicz-Nowikow (2014) discussed how gaze is being utilized as a pointing gesture. The study concluded that Polish people tend to point at objects with their index finger when necessary but find it rude to point at people. They would rather point with an open palm or even with just gaze when pointing is necessary.

## 2.2. Conversation analysis as a method

CA is utilized as a method mainly in sociology, and Harvey Sacks, an American sociologist, is seen as the method's founder. He started the first lecture of the series in 1964 without giving any background information and went headlong explaining his analysis. The tone of research was set then: first gather the data, look at what you see and then do the analysis. Sacks' series of lectures began with him giving quotations from a telephone conversation collected at an emergency psychiatric hospital. After presenting the excerpts, he moved to discussing the "rules of conversational sequence" and compared the conventions of greetings in audio recordings. He was interested in the types of responses the different ways of initiating the greeting sequence would produce. (Sacks, Jefferson & Schegloff 1995, pp. 3-4) Already during the first lecture Sacks was able to exemplify the principle of CA research: data defines what is to be researched.

Even though the principle of CA is apprehensible, the exact method of CA is impossible to determine. Jack Sidnell (2010) says that "at its core, conversation analysis is a set of *methods* for working with audio and video recordings of talk and social interaction" (p. 20). The problem with trying to depict how CA is conducted is that there are no written rules on how to do it exactly (ten Have 2007, p. 11; Sidnell 2010, p. 20), because its fundamental goal is to reveal things no one imagined there being. However, there are rules that must be followed by a researcher in order for a study to be comprehensible for others.

First, the data must be transcribed to a written form. According to Sidnell (2010), in CA the transcribing process is not the obligatory boring part of research. It is the process during which the researchers familiarize themselves with the data and try to hear what was said. Transcribing is a crucial part of the method and analysis. The process helps with developing an ear for discovering nuances and finding patterns in speech. Hearing something and marking it as exactly what was said is a skill that can only be learned by repetition. Furthermore, repetition makes the process itself faster and easier, which helps when doing further research.

Second, the transcribing process is usually the point where the researcher will discover whether there is anything worth of studying in the data or not. CA is not the method for theorizing what people meant by doing something. That is to say, Goffman (1963) says that conversation analysts should try to “identify the basic substantive units, the recurrent structures and their attendant processes” and they should try to answer “what sorts of animals are to be found in the interactional zoo”(Goffman, 1963, as cited in Sidnell, 2010, pp. 28-29). The data provided by CA desires to be unbiased and purely informative. When doing the analysis researchers should focus on describing the data and the interactions in it instead of trying to overanalyze subjects’ actions and coming to conclusions that are speculative in nature. In CA researchers should not try to read the minds of their subjects. However, even though researchers should not speculate too much on participants’ interactions, categorizing different actions calls for researchers’ knowledge of the target language and the “gut feeling” the researchers possess for example the fact that structure “how are you doing” is in most of the cases a question does not have to be explained through grammar of the English language.

There are different transcription systems, but most of them tend to follow the rules created by Harvey Sacks in the early years of CA. The transcription style used in this study is loaned from Jefferson (2004). It is a clear and consistent style that is easy to follow, after one has familiarized oneself with it. A concise version of the transcription conventions is quoted from Enfield and Sidnell (2017) (see Appendix A). In addition to the conventions of Enfield and Sidnell I will be using pound sign (£) to indicate that the speaker is smiling while speaking. The multimodal transcription system is loaned from Mondada (2007) (see Appendix B).

After collecting the data and transcribing parts of it one should start analyzing it. In Sidnell (2010) it is suggested that one should start by looking for patterns. There are four ways of approaching the pattern gathering process. First, one can look for patterns across data samples, which occurs when

one has gone through many different recordings and finds out that they have seen something similar before and gathers the data from different recordings. Second, one can find the patterns within one set of data for instance repetition next to certain utterances or a recurring term. Third, instead of focusing on the forms used one can “track the various forms selected to refer to or formulate some referent” (p. 30). Fourth, one can observe the types of formats for doing an action. As an example, speakers may “be seen to select among different formats in other-initiating repair, choosing between, for instance, ‘what?’ and ‘you went there?’ and ‘you went to the *park*?’ and so on” (p. 30). After one of the listed types of phenomenon has been chosen, a collection of should be gathered instead of only looking at one instance. A collection consists of multiple instances of the same phenomenon, because only then different types of aspects or features can be revealed and then “we can see the range of actions a given practice can implement” (pp. 31-33).

Final step in the whole process is developing the analysis. Sidnell (2010) says that “there are many ways to do” the organizing of data and he offers a suggestion for the matter. The way the analysis is conducted in this study is based on his suggestions and it is as follows:

1. Cases of the phenomenon of interest are gathered and transcribed.
2. Then the transcriptions will be sorted into subsets of the same phenomenon.
3. All the cases will be discussed separately under their own subheadings and only necessary parts of the transcriptions, not the entire transcriptions, will be presented in the text.
4. Finally, the phenomenon be discussed in section 5 (p. 34).

### 3. Data

Section “Data” is divided into two subsections: “Data” and “Gathering data from an institutional setting”. Topics of data gathering, transcription method and initial findings will be discussed first, and later the topic of institutional setting will be discussed. It will be argued why the data chosen for this study is fitting for CA research and why the setting of debate is natural.

#### 3.1. Data

The data presented and analyzed in this study are excerpts from two different video recordings. The recordings are from two debates: South Carolina Republican Debate recorded by CBS in early 2016 and a Nevada Democratic Debate recorded by CNN in late 2015. The participants of the Republican Debate are debaters Ben Carson, Marco Rubio, Donald Trump, Ted Cruz, Jeb Bush and John Kasich. Then there are John Dickerson, Major Garrett and Kimberley Strassel, whom are referred to as Judges in the excerpts. The participants of the Nevada Democratic Debate are the debaters Lincoln Chafee, Hillary Clinton, Martin O’Malley, Bernie Sanders and Jim Webb. The judges of the Democratic debate are Anderson Cooper, Dana Bash, Don Lemon and Juan Carlos López. The reason for referring the interviewers as Judges is to avoid confusion; it is immediately clear to the reader which one of the participants is the debater and the judge. Audience will occasionally clap, yell or some other way intervene with the participants and they shall be referred to as Audience in the excerpts.

The data will be transcribed manually for multiple reasons; however, automated transcriptions have been discussed as an option for the laborious manual work of transcribing conversation. Even though Moore (2015) admits that automatically produced transcripts should not be used as final versions of transcriptions, he sees opportunities in these automatically produced transcripts. For example, some automatically produced transcriptions have the error rate as low as 2% which is close to the first transcriptions done by hand (pp. 268-269). Therefore, using machines instead starting from the scratch might save a lot of time in some cases. The time-saving perspective becomes even more important when transcribing conversation corpora of immense volume. Due to the modern



digital age accessing data has become relatively easy. Therefore, having the ability to get instant access to even partial transcriptions of new data would benefit analysts. However, there are problems with automated transcriptions.

Bolden (2015) expresses their concern by listing benefits of manual transcription and disadvantages of automated transcription. The process of transcribing is a key element of conversation analytic research and training. Analysts are supposed to have an intimate relationship with their data, where close and careful listening takes place. Transcribing from scratch makes researchers analysts rather than observers of interaction. In addition, analysts can discover new phenomena which were inaccessible to automated transcription. Furthermore, transcribing is an important part of CA training, because the best way for new analysts to learn transcribing is starting from the basics of it.

Another problem with using automated transcription is that they are only able to transcribe high quality recordings in which the participants speak in standard English. This would warp the integrity of the data collections to undesirable directions as it is unwise to research only speakers of one English dialect. Lastly, the labor-saving aspect falls flat in practice. Automated transcriptions work adequately when only one participant is speaking at a time, but when multiple participants are speaking over each other, the system is not able to differentiate them from each other. Therefore, analysts will have to go through every transcription and correct the mistakes made by the machine and suddenly no time had been saved (Bolden, pp. 276-279). Thus, in this study the data was transcribed by hand. Even though there was not too much of participants speaking over each other, using automated transcription would have not given the same benefit when looking for the points of interest in the data.

After listening the debates, I was instantly interested in the way the debaters were answering questions, because sometimes the debaters seemed to be more interested in avoiding the question than answering them. The turns that raised my interest were the answers that started with either *well*, *well first of all* or *first of all*. After those three utterances the answers seemed to go off trail more often than during answers starting with something else. If the debaters did not answer the question straightforwardly after the utterance, they would either do a sort of a circumbendibus which would eventually go back to the original question or the debaters would not answer the question at all. Following Sidnell's (2010) guidelines of gathering recurring examples, I gathered a collection of the types of answers to a question that began with any of those three utterances and divided them into four subsets: the addition, the "not answering the question" (NATQ), the

“commenting on a prior question” (CPQ) and for comparison I counted the straightforward answers, as well. Naturally, there were other instances where the debaters began their turns with either *well*, *well first of all* or *first of all*. However, during those instances the debaters were not always answering questions. They were either rebutting another debater’s argument, answering a question presented to every debater in turn, answering a pursuit or giving a list of some sort. I decided to delimit my research topic to debaters answering question appointed to a specifically them and therefore the other instances would not be included in the analysis. After counting the instances of subsets and straightforward answers emerging in the data the following graph was produced:

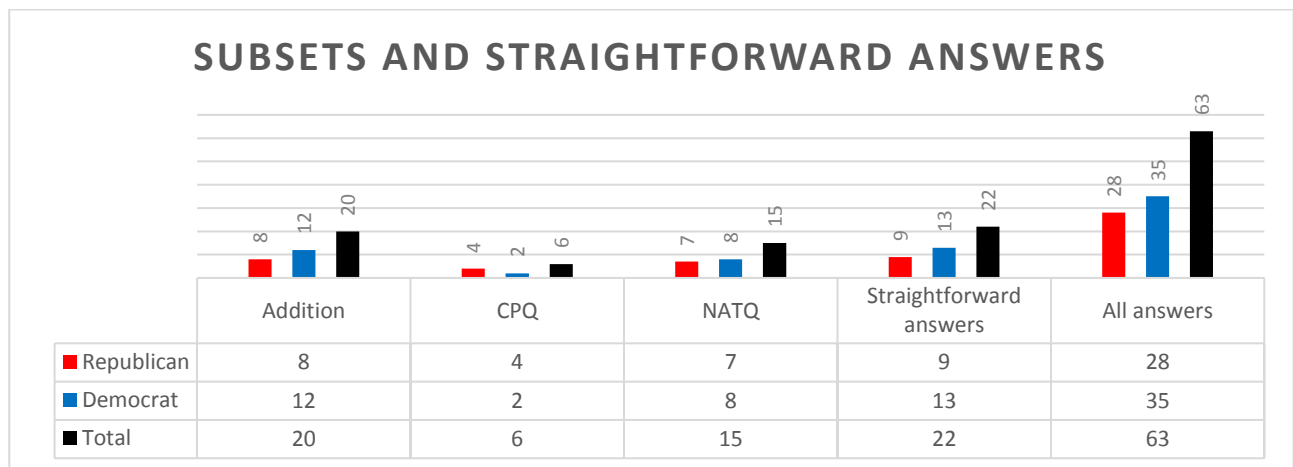


Figure 1

In the Republican debate there were 28 answers in total and in the Democratic debate there were 35 answers in total, which makes the total amount of answers beginning with *well/first of all*, including straightforward and non-straightforward, in both debates 63. Only a bit over third (22) of the 63 answers were straightforward answers. This fact made me delimit the topic even further and I decided to focus on the non-straightforward answers, because politicians are often accused of diverting questions and trying to stay as neutral as possible in order to not lose voters. Therefore, it would be interesting to focus on those turns where politicians are not answering the questions right away but are trying to discuss something else first. It is worth noting that those types of answers are not the minority but the majority of *well/first of all*-prefaced answers. Most of those non-straightforward answers were additions or NATQs and there were 6 CPQs, which shows that most of the time the debaters wanted to add something before they would answer a question, or they would rather not answer the question at all. Commenting on a prior question was a rarer occasion, and it raises a question if it should have been a subset of its own at all, but its justification will be argued for later when I will be discussing the subsets further in section 4.

### 3.2. Gathering data in an institutional setting

There are some rules when it comes to gathering data in CA. The data has to be accessible and it is most preferably a video or audio recording of interaction in an authentic environment. What is meant by authentic is that the interaction should take place in a non-laboratory setting. The recording space should be a place where the participants have either been before or they feel comfortable or in the best-case scenario both. In addition, the participants should feel as little as possible like they are being monitored during the recording by an observer. Therefore, there should be only as few recording equipment as possible 'lying' around the recording space. Moreover, the researchers should not be around while the recording takes place for the same specific reason. Any sort of disturbance of the recording should be avoided.

Initially, it would seem like ensuring that there are as few distractions as possible would be enough. However, it has been proven that participants might change their way of speaking to a more formal one, because they know they are being monitored by a group of researchers. Labov (1972) came to the realization that when prior scholars researched phonological structures of communities in a big city, they found that "the speech of most individuals did not form a coherent and rational system, but was marked by numerous oscillations, contradictions and alternations." However, when speech of individuals "was charted against the overall pattern of social and stylistic variation of the community, his linguistic behavior was seen to be highly determined and highly structured." (pp. 123-124) It became evident that even if researchers tried to take every variable into account, they could not record speech that would not be affected by their preparations or the setting they had produced. Thus, the term natural has been challenged by critiques of CA.

It has been argued that the setting is not natural even if all the factors presented in a prior paragraph are considered. According to ten Have (2007) it has been argued by the critiques that researchers using CA are victims of so-called *observer's paradox*. It has been said that the method does not produce naturally occurring data, because "the mere recording of the data disrupts and transforms it" (p. 34). However, Labov (1972) acknowledged the paradox and suggested that it "is of course not insoluble: we must either find ways of supplementing the formal interviews with other data, or change the structure of the interview situation by one means or another" (p. 209). Labov goes further as he suggests solutions such as using devices "which divert attention away from speech"

(p. 209) that could lead to more free expression of vernacular speech by the actor. Those kinds of devices have been used in recent studies such as when in a study by Mondada (2014b) the participants were placed in a situation where a professional chef was teaching and instructing them on how to prepare a dinner. The participants were concentrating so much on the task at hand that they would not be distracted by the cameras. Therefore, Labov's suggested strategy has been utilized in research, however, the fact that it has been utilized does not in itself prove that natural data can be produced.

To further prove that gathering natural data is possible, it has been argued that the word natural refers to the idea before collecting the data (ten Have, p. 35). The action recorded should be something that may have taken place without the researchers' effort. Therefore, CA is perhaps the best method if one desires to research something naturally occurring, because the activities in question should not depend on the researchers' intent; researcher is there just to observe what phenomena are taking place during interaction. Furthermore, arguing whether anything would take place without some interference by an external source of influence is in itself a topic that has no clear solution to it with the tools of linguistics research we possess. Discussing that topic is highly philosophical and should be left for the scholars of those fields for now. I would argue that discussing whether results are valid is reasonable on a level of individual studies, and criticizing methodologies is reasonable, too. However, playing word games and discussing what natural really means is irrelevant. As long as the possible interference is taken into account, discussed and dealt with the natural has been achieved.

Labov (1972) discusses the problem of gathering data from mass media. He thinks that it is "possible to obtain some systematic data from radio and television broadcasts," but at the same time he thinks that "the style is even more formal than that we would obtain in a face-to-face interview." Thus, the data would not represent the general population. (p. 211) I agree that politicians' behavior would be the same when they are discussing different political topics with for example their relatives. Nevertheless, the debate environment is a natural setting of its own with its own rules and agreements and the participants' roles are extensions of their "natural selves" accommodated to the situation at hand. We all accommodate our behavior depending on the situation at hand even when we are not acting in an institutional setting. Dinner with friends and dinner with family may differ from each other and have different rules. Thus, it is the rules of the setting that guides actors' in certain situation.

The participants in debate situations follow the rules of the setting. Usually a judge presents a question or hypothetical situation to which he asks an answer or comment from one of the debaters after which the judge may ask further questions or give another debater an opportunity to rebut. For the participants the format of debate is familiar especially because they have teams that have prepared them for a debate. Debate is a natural setting for politicians. They know the rules and they certainly know when they are breaking them. Therefore, it is not only the participants acting the way they see fitting, but the rules of interaction in the debate setting create expectations: the participants know how the audience, judge and the other participants assume them to act. Some might say that the multitude of cameras would compromise the naturalness of the setting, but I argue that the cameras filming the debate and debaters acknowledging their presence are both part of the setting of a public debate and thus they do not compromise the definition of a natural setting. The participants just modify their behavior according to the new natural setting they are put into.

As can be seen from Sacks, Jefferson & Schegloff's (1995) research back in the beginning of CA, institutional encounters were the first interactions to be analyzed utilizing CA. Peräkylä (2013) discusses institutional encounters, in his example psychotherapy, and how particular inferential frameworks are to be seen in such interactions. When in "normal" interaction people try to understand the other participant's means of communication as them trying to display their communicative intentions, in psychotherapy therapists try to "examine the patient's talk beyond its intended meaning" (pp. 552-553). Political debates are similar in the sense that some of the listeners of the debate try to examine the politicians talk beyond its intended meaning, however, debates are different in the sense that some people in the audience are only in it for the ride. Furthermore, to be certain that debates are institutional talk, one has to examine the interaction further and compare it with the three basic elements of institutional talk listed here (Drew & Heritage, 1992, as cited in Fitch & Sanders, 2005):

1. The interaction normally involves the participants in specific goal orientations that are tied to their institution-relevant identities: doctor and patient, teacher and student, bride and groom, and so on.
2. The interaction involves special constraints on what will be treated as allowable contributions to the business at hand.
3. The interaction is associated with inferential frameworks and procedures that are particular to specific institutional contexts (p. 106).

Debaters have the goal of gaining more voters on their side than the other debaters. Voters are the ones in control of who will get the votes, and their goal is to figure out which of the candidates would advocate their opinions if elected. These two participants are tied to their identities of *debater* and *voter* (hence listener). The third role in the debate belongs to the *judge*, who acts as a communicator between the audience and the debater. Their role is to ask questions to which they think the debaters should answer to. They also have to ask the question in a manner that gets a proper answer out of the debater. All the institution-relevant identities in debate are constant and do not change during debates. Thus, the first element is present in debate setting. The second element is also present, because rules of the debate bind the participants to their roles. The audience can yell, applaud, cheer or boo, but it is the judge who asks the questions. Debaters are allowed to answer the questions and attack other debaters, but they are not allowed to have a two-way conversation with the audience as the audience cannot reply. Finally, the third element manifests itself in the rules of the debate, which are specific to the debate setting and often agreed upon in the beginning of the debate. Debate being part of institutional talk is further proof of it being a natural setting, thus, it is a relevant topic to study in CA.

## 4. Analysis

When gathering the data, it became obvious that while sometimes when the participants used the *well/first of all* they answered the question right away, most of the time they tended to say something else first instead, if they even answered the question at all. I was able to categorize the non-straightforward answers to three subsets of answers: ‘the addition’, ‘not answering the question’ or NATQ, and ‘commenting on a prior question’ or CPQ. The three subsets will be discussed under their own subheadings. Each subheading will begin with a short introduction and followed by three descriptive excerpts of that subset. Subheading will end with a short summary that will be discussed further in section 5.

Before getting into the analysis it is important to acknowledge that the debaters knew the rules of the debate. The rules were set at the beginning of the Republican debate when it was said that: “When we (the judges) ask a question you will have one minute to answer and thirty seconds more if we ask a follow-up. If you are attacked by another candidate, you get thirty seconds to respond.” The rules were exactly the same in the Democratic debate. Therefore, the rules suggest that the debaters knew they had limited time to answer and they chose to answer accordingly. It was never a matter of them not having an opportunity to answer or them not knowing how much they had time to answer. Now to the first subset the addition.

### 4.1. ‘Addition’

The first subset discussed is the addition. It is defined as follows: the addition is a *well/first of all*-prefaced expansion after the FPP of an adjacency pair which gives additional information that does not seem to relate to answer itself but may or may not contribute to the answer in the end. Sidnell (2010) mentioned that conforming the adjacency pair is not a probability but “rather, the organization described is a norm to which participants hold one another accountable” (p. 64). The addition will prove that the pair-conforming response is often the expected but never a certainty.

The addition is closely related to the concept called topic shift that Warren (2006) describes as an instance “when there is no apparent connection between successive utterances in relation to the existing topic framework and which in effect mark the creation of a new topic framework” (p. 173). However, the difference between the addition and topic shift is that where a topic shift tries to initiate a new topic, the addition works as either a gateway to a topic shift or a topic drift, or as a circumbendibus to the original topic. An example of the latter can be perceived from the first excerpt after some background information.

The first excerpt is from the Republican debate. Judge is discussing what Carson became known for during early stages of his campaign. One of the things he used for gaining popularity was him having enough of the political correctness that, according to Carson, had become the norm in American politics. Before the excerpt Judge even claimed that it sparked his rise during the campaign. Then Judge proceeded to request Carson to tell something to the audience that is of importance but is considered politically incorrect. Consider the following excerpt:

(2) *Dr Carson I*

1 JUDGE: politicians are often accused of (0.6) g(h)lossing any hard  
 2 choices people have to make. .hh just always selling <happy  
 3 nice things>. (0.6) ↑so (0.5)in the uh ↓in spirit of saying  
 4 something that might be politically incorrect, (.) tell the  
 5 voters something that they ↑need to ↓hear but that ↑might be  
 6 politically in↓correct.  
 7 (0.4)

8 CARSON: .hhh (0.5) for first of all (.) I’m not a politician, (0.4) so  
 9 i’m never gonna become a politician .hh (0.4)  
 10 \*but here’s what ((swallows)) >here’s what people need to know<,  
 11 *car* \*hands are in palms against each other position and are moved  
 12 back and forth--->  
 13 (1.0)

14 CARSON: #people need to know (0.5) that free college, (.) is no-, (0.6)  
 15 *fig* #Fig. 2a

16 CARSON: >it is# a< non-starter\*.  
 17 *car* --->\*

18 *fig* #Fig. 2b



Participants Judge and Carson performed the directive/response adjacency pair and not the question/answer adjacency pair, because Judge's last sentence on line 4 began with an imperative form *tell* and Judge was asking Carson to tell it *to the voters* instead of asking his opinion on the subject. What is expected from a debate is often questions and answers, but something that was prominent during this Republican debate was that the judges were eager to direct the debaters to tell something to the audience instead of asking questions from them. Unlike in the study of Etelämäki & Couper-Kuhlen (2017) the target of the directive does not resist.

After Judge made the directive, Carson's response started with an insert expansion on line 8 *for first of all I'm not a politician, so I'm never gonna become a politician*. Normally, insert expansions can be divided into either post-firsts or pre-seconds (Schegloff, 2007, as cited in Sidnell, 2010, p. 103). The expansion in question falls under the post-first category, because it addresses the Judge's first sentence on line 1: *politicians are often ... happy nice things*. However, the insert expansion is not related to the Judge's directive itself. Carson was adding a comment related to Judge suggesting on lines 1 and 2 that *politicians are accused of glossing any hard choices*. When Judge said *politicians are often accused of* Carson excludes himself from the group of politicians by saying he is not a politician. Beginning on line 8 he insisted to the audience, not to Judge, that he was not one of politicians presented by Judge: *I'm not a politician. So, I'm never gonna become a politician*.

It could seem like the addition is no more than an ordinary topic shift: Carson introduced a new topic after Judge had finished his turn. The window for topic shift might have been open from a turn-taking point of view, but because Judge presented a directive, the anticipated response was a response as in directive/response adjacency pair. Carson did not respond in a straightforward manner, because first he performed the addition instead. However, instead of keeping on discussing the new topic, Carson returned to the original topic after the inbreath on line 9. Usually, when a topic shift occurs, there would be some interaction between the participants regarding the new topic; Judge could have commented on Carson's new topic of being a politician. However, in this case Carson had no intention to stay in the new topic for long, which can be concluded from the fact that it is him who starts discussing the original topic after his comments. That is the difference between a topic shift and an addition.

After the addition, Carson moved on. The repairs that took place on lines 10, 14 and 16, when he carried on with his response, suggest that the addition was a needless supplement: *but here's what, here's what people need to know, people need to know*. There was an absence of topical coherence

between what Judge said before the addition and the addition, and between the addition and the pre-second expansion. Carson repaired twice during the pre-second expansion before his response of the directive/response adjacency pair. He could not go straight to his response from the addition, because of the absence of topical coherence. He had to mark the shift with an expansion, which in this case was accompanied by multiple repairs. Marking a shift is not limited to just the addition as it will be evident later on with the CPQ, too.

The response is pair-conforming, because Judge asked Carson to tell *the audience something that they need to hear but that might be politically incorrect*, and Carson said that free college is a non-starter. Even though delivering the response to the directive was postponed due to the addition, Carson was able to respond pair-conformingly eventually. However, whether the answer was insufficient is a more difficult topic. After Carson said that free college would be a non-starter, he does not present further proof why it would be impossible to implement. Beyond the excerpt, he began to discuss the current difficult economic situation and how the Democrats are planning to tax money out of rich people when the problem is, by his own words, *the irresponsible evil government*. It is true that free college would depend on government/state funding and that a tough economic situation would require either reorganization of government/state funds or redistribution of wealth amongst the population via taxation, but neither of those ideas are impossible in themselves. However, Judge did not demand for a more nuanced answer from Carson after he had finished his turn. Therefore, as a conversation analyst we have to trust on Judge's judgement on debater's answer being sufficient enough for him not to answer further questions. The response was therefore sufficient.

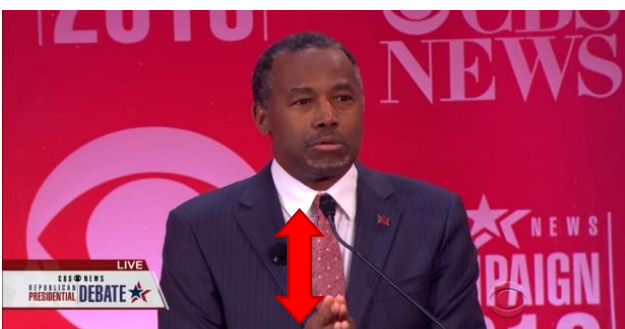


Figure 2a



Figure 2b

According to Mondada (2014a), seeing interaction as fundamentally multimodal enables to focus on the interplay of multiple resources. Therefore, all the excerpts will be analyzed from the perspective presented by Mondada (2014a) and it is preferable to find something multimodal in

them through thorough examination. After Carson claimed he is *never gonna be a politician* on line 9 he addressed the question from line 10 to 16. Before his answer, he moved his hands in front of him in a prayer-like position as can be seen in figure 2a and began to use his hand as a sagittal beat gesture during his answer. Every beat of his hands took place during a different word. In Krahmer & Swerts' (2007) research the findings on visual beats (nod, hand movement, raising eyebrows) suggested that just before uttering a word beat gestures are often used as part of the Gestalt by the speaker as to mark the important parts of speech. In addition, the listener pays more attention to the highlighted parts as was concluded in Biau & Soto-Faraco (2013) where they found that the speaker might perform the gesture a bit earlier than the word to emphasize it and the listener interprets the following utterance as something of higher importance. In addition, they concluded that a speaker may use the gesture to parse their speech, which is also true in Carson's case as at *is not, it is a* on lines 14 and 16 he closed his eyes as can be seen from figure 2b and performed beat gestures during the repair. Another interesting detail in the use of gestures is that Carson utilizes a gesture during the shift from the addition back to the original topic. The occurrence of gestures within a joint of two topics is not unique to the excerpt 2 but will be present in later excerpts, too.

In the excerpt 2 the response is pair-conforming and sufficient. However, whether the response was completely sufficient or not is subjective. The marker we are using as to define the fact is if the Judge asks for clarifying comments or pursues the answer further. A pursuit must be one of the four types listed by Romaniuk (2013): the explicit reference to the initial question, the verbatim repeat of the question, the indexically linked pursuit or the implicitly formulated pursuit. In the next excerpt the debater addresses the question, explains his views back in the time of him doing the things he did and explains his views now. In addition, Judge does not ask a follow-up of any sorts.

Before the following excerpt from the Democratic debate, Judge was discussing the candidates' qualifications for becoming the commander-in-chief of the United States. After discussing with Jim Webb, a former marine, Judge turned his attention to Bernie Sanders. The choice created contrast between the candidates, because he just questioned a former marine and Sanders was a pacifist in his youth and applied for conscientious objector status during the Vietnam War. Sanders' application for conscientious objector status during the Vietnam War was one of the major targets of critique when it came to his campaign and many questioned his adequateness for becoming the commander-in-chief. Judge decided to question Sanders about his past views. Consider the following:

(3) *Conscientious objector status*

1 JUDGE: senator sanders tell an american soldier who's watching right  
 2 now tonight in afghanistan why you could be commander in chief,  
 3 (0.4) given that you app[plied for conscientious] >objector  
 4 SANDERS: [Well for \*first of all]  
 5 san \*.....  
 6 JUDGE: status.<  
 7 (0.2)  
 8 SANDERS: let# \*me: applaud#\*, (.) my good friend\* jim webb, (0.5) for  
 9 san ....\*points-----\*,,,,,,,,,,,,,,\*  
 10 fig #Fig. 3a #Fig. 3b  
 11 SANDERS: his service to this [country in so many ways. (1.3) uhm (0.8)  
 12 AUDIENCE: [audience clapping]  
 13 SANDERS: jim and i:\*,\* (1.2) under \*jim's#\* leader\*ship as he indicated]  
 14 san \*.....\*points\*,,,,,,\*  
 15 fig #Fig. 4a #Fig. 4b  
 16 SANDERS: passed the most significant, (0.6) veterans' education.(0.2)  
 17 bill. (0.8) uh in recent history. (0.7) uh we followed suite,  
 18 (.) with, (0.3) few years later passing under my leadership.  
 19 (0.6) uh the most significant.= veterans', (.) health care  
 20 legislation. (0.6) uh in the modern history of this country.  
 21 (1.0) [I:(1.3) when i] was a young man,  
 22 AUDIENCE: [audience clapping]

Judge and Sanders perform the directive/response adjacency pair as Judge and Carson did in the excerpt 2. The directive is formulated similar to an order as was in excerpt 2, too. After pronouncing the title and name of the debater Judge begins his directive with imperative *tell* on line 1, and instead of requesting Sanders to tell him, he directed Sanders to tell it to an American soldier *who's watching right now tonight in Afghanistan*. This is similar to the form of the request in excerpt 2.

Before the directive was finished by Judge from line 1 to 5, Sanders had already started to perform his response during *applied for conscientious* on line 3. The post-first insert expansion from line 4 to line 8 *well first of all...many ways* is related to the discussion before the directive. Judge had discussed Webb's expertise as a military leader with Webb and it became clear during the conversation that Webb was a member of a family that had a long history of joining the military. From line 8 to 10 Sanders utilized the opportunity to applaud the opposing candidate Webb for his

service for the country, which makes the incident to belong to the category of the addition. The additional information seems to not relate to the request presented nor to the response to follow the additional information. Sanders pulls Webb as an object to his addition, which works as a transition to the beginning of Sanders response.

Sanders begins his response from line 13 by telling the audience that him and Webb, a marine veteran, have worked together for the better of veterans. According to Sanders from line 13 to 20, they have passed a *veterans' education bill* and *the most significant veterans' health care legislation* together which is an answer to the veterans why Sanders could be the commander-in-chief: he cares of their military and their veterans. From line 21 onwards Sanders addresses rest of the question which was *given that you applied to conscientious objector status*. He starts from him being a young man back then, that can be seen in the excerpt, and ends with telling the audience that he is no longer a pacifist, as he has supported wars in Afghanistan and Syria. However, he believes that war should be used as the last resort and he would rather exercise diplomacy first. The response was created in such detail and all the answers were related to the directive that it is fair to say that the response was sufficient. Another indicator of sufficiency was that Judge did not ask any further questions related to the topic.

Sander's phrase *well, first of all* on line 4 functions the same way as Carson's *for first of all* in excerpt 2. It marks a topic shift in the conversation he is about to perform, even though the expected way to start the response would have been something directly related to the directive. In Carson's case in the excerpt 2 the addition was completely unrelated to the directive and it was therefore easy to determine as a case of the addition. However, Sanders' case brings up the most challenging problem faced when determining the addition throughout the data.

The addition in excerpt 3 that followed the possible topic shift seemed to be unrelated to the response initially but was to reveal itself to be related to the following argument. Therefore, it has to be determined whether a topic shift occurred or not, because if it did not, there was no addition. Starting with looking at the phrase *well, first of all* alone suggests that either Sanders was about to present a list of multiple arguments for his cause or he was more interested in discussing something else first. Telling the audience that Jim Webb has achieved great things in the military is not an argument for Sanders being a possibly great leader. However, in Sander's case the function of the phrase was to emphasize the following argument of him working with Jim in the past. On one hand applauding Jim Webb is not an argument for Sanders, but on the other hand it is somewhat linked

to the points made by Sanders after the applauding. A ruling for all the data had to be made which was: if the *well/first of all*-prefaced response does not initially seem to be related to the question/directive, it is a case of the addition. As mentioned earlier, Sanders' response is sufficient. The response was argued from all viewpoints and Judge does not ask further questions from Sanders. In addition, the response was pair-conforming as Sanders performed what was expected from him. The only issue with the answer was that it was anything but straightforward. Instead of going straight ahead to the legislation he had prepared and passed in the past, he pulled the other candidate, a marine veteran, to the discussion and piggybacked on Webb's career as a marine and an employee in the Pentagon. Thus, even though the answer was pair-conforming and sufficient, it was lacking straightforwardness as was Carson's answer in excerpt 2.



Figure 3a



Figure 3b

Sanders' action has multimodal properties as it includes distinct gesticulation related to his speech. During Judge's directive from line 1 to 6, Sanders' had already started to prepare his action by raising his hand on line 5. The preparation is accompanied the words *first of all let me* on lines 4 and 8, and during the preparation Sanders raises his right arm and extends his fingers into a pointing gesture as in the figure 3a. The preparation leads to the action, where Sanders points twice at Webb. First, he points during the word *me* and then he points during the word *applaud* as can be seen in the figure 3b, and after a short pause the pointing gestures are followed by the words *my good friend Jim Webb for his service to this country in so many ways* from line 8 to 11 and the retraction of the gesture. As was mentioned earlier, after the words *well, for first of all* Sanders shifted the focus to his co-debater. This observation is supported by the visuospatial modality, because Sanders pointing at the person he is talking about. This suggests that Sanders would like the listener's attention to be somewhat shifted to Jim Webb. He is shifting the focus from himself to the other debater, thus, speaking of Webb's achievements does not seem out of place.



Figure 4a



Figure 4b

After speaking of Jim's achievements, Sanders moves to his response by saying *Jim and I* on line 13 which is followed by another preparation for a different pointing gesture as can be seen in the figure 4a. The gesture apexes during the word *Jim's* on line 13 and is followed by a gesture's retraction right after as can be seen in the figure 4b. This pointing gesture is a brief flick of hand towards the person Sanders is referring to. All three pointing gestures were performed when Sanders was talking about his co-debater Webb to whom he was referring to.

Sanders made sure that the audience would notice Webb during the response, and thus he was able to bring Webb into the discussion. When Sanders said *for first of all* on line 4, a topic shift took place. The new topic of discussion was Jim Webb and how Sanders appreciated his service to the United States. After Sanders had performed the addition, another topic shift occurred which took the discussion back to the original directive. Interestingly, Sanders' pointing gestures took place during the two topic shifts before and after the addition. The first one was before addition on line 4 and the second one was after the addition on line 13. The way the addition was organized made it possible that after praising Webb for his service to his country, Sanders was able to tag him to his answer quite easily. In his response he was now able to point out how Jim, a marine veteran, and him had cooperated in the past by passing pro-veteran legislation. Sanders used Jim as a way to shore up his legitimacy as a commander-in-chief. One could say, in CA terms, that even though the object "Jim Webb" does not provide instantly perceivable affordances to anyone, with the right combination of gestures and speech Sanders was able to utilize Webb's experience as a marine veteran and Webb's answer to the question addressed to him before Sanders' turn. Sanders telling the audience that he has worked on multiple occasions with the veteran, who just said he is the most qualified for becoming the commander-in-chief, makes it sound that even if Sanders is not the best option, he is not too far off.

In excerpt 3 the answer was sufficient, because Sanders told why he would be a good commander-in-chief despite his past and Judge did not ask him further questions related to the topic. However, the response was non-straightforward as was Carson's response in excerpt 2, because Sanders did not go right into the answer and decided to discuss something else first. Furthermore, visuospatial modality was utilized in the points between topic shifts as was in excerpt 2, too. In the excerpts 2 and 3 the addition served as an addition of information before the response was performed. It may or may not be related to the directive, but nevertheless its goal was to add something before carrying on to the response. However, in some cases the addition served a different function as can be seen next.

In the following excerpt from the Republican debate Judge was quoting Trump's prior statements about the Iraq War. Trump had said that George Bush should have been impeached for getting the United States in to the war and that impeaching Bush would have been a wonderful thing. The problem is that the consensus within the Republican party was and has been that Bush did what had to be done with the information they had when he with his cabinet made the call to invade Iraq. They were in the knowledge that Hussein might have had obtained weapons of mass destruction and they had to make a decision. A reasonable explanation for Trump's prior statement is that before joining the Republicans Trump had been a supporter of the opposing Democratic party. Naturally, when Trump became a candidate for the Republican Party this had become a question of interest. Many had wondered would Trump take back his statement or would he stay in his position. Consider the following excerpt:

#### (4) *Impeaching George W. Bush*

1 JUDGE: you said QUOTE, (.) which personally i think would've been a  
 2 wonderful thing,  
 3 (0.6)  
 4 uh close quote, when you were asked what you meant by that you  
 5 said for the war , (0.4) for the war he lied he got us into the  
 6 war with lies.  
 7 (0.5)  
 8 do you still believe president bush should be impeach-  
 9 >should've been [impeached?<]  
 10 TRUMP: [first of all \*i] have to say as a businessman  
 11 tru \*.....



12 TRUMP: i get# along \*with everybody. i've business ↑all\* over\* the  
 13 tru .....\*two sagittal sweeps with his hand\*,\*,\*,\*,\*  
 14 fig #Fig. 5a  
 15 TRUMP: ↓world,  
 16 ((yelling and occasional clapping in the audience for 3.9s and  
 17 something hitting the mic twice during the 3.9s))  
 18 TRUMP: i know \*so many of the people in\* the# audience. <and >BY  
 19 tru \*points with right index finger\*  
 20 fig #Fig. 5b  
 21 TRUMP: THE WAY< I'M A SELF-FUNDER, (.) I DON'T HAVE I HAVE MY WIFE and  
 22 i have my son. (0.5) THAT'S ALL I HAVE, (.) I DON'T have  
 23 °(those).°  
 24 (0.9)  
 25 TRUMP: \*#SO, (1.1) let me just tell you.\* (1.0) i get along with  
 26 tru \*raises his hand and makes a baton gesture\*  
 27 fig #Fig. 5c  
 28 TRUMP: everybody which is my obligation to my company °to myself et  
 29 cetera.° (0.6) obviously (.) the war in iraq was a big fat  
 30 mistake, all right? (0.5) now you can take it any way you want,  
 31 and it TOOK JE- it took JEB BUSH,

In excerpt 4 the adjacency pair in question is question/answer, which is usual for debates and possibly even more usual than a directive/response. Judge asked Trump if he still was in the opinion that president Bush should have been impeached for, paraphrasing Trump's own words, getting the United States in to the Iraq war with lies. To a polar question the pair-conforming straightforward answer could start with something simple as either "yes, he should have been impeached" or "no, he should not have been impeached", and then the arguments for the answer could follow. However, Trump started his answer with the addition starting on line 10: *First of all, I have to say as a businessman I get along with everybody*.

The addition starts on line 10 and acts as an expansion, that has three expansions within itself, between the answer/question adjacency pair. It is impossible to define whether this particular expansion is a post-first or pre-second expansion, because Trump did not respond to anything Judge said before nor are the expansions related to Trump's answer. What is certain is that because Trump said *for first of all, I have to say* on line 10, there is something he wanted to say something before

answering the question presented. He wanted to provide additional information; thus, this is a case of the addition.

The addition consists of three different expansions: first Trump claims *I get along with everybody and have business all over the world* from line 12 to 15. Second, he told that he is a *self-funder* supported only by his family and has no lobbyists on his side from line 18 to 23. Third, he repeated that he *gets along with everybody* from line 25 to 28. Then Trump finally moved to the answer which started by him saying from line 29 to 30 that the *war was a big fat mistake*. The reason why the addition took so long was that starting on line 16, between the first and the second expansion, the audience reacted to Trump's claims by cheering and applauding for 3.9 seconds. Worth noting is that none of the three expansions were related to each other, but they will be defined as a single addition, because the three expansions as a whole was what Trump wanted to say before he moved to the answer.

There was no coherence between the question and the addition nor between the answer and the addition. The question was not about Trump's personal relations or how he funded his campaign and it was not necessary to address these subjects for the sake of the answer either. Therefore, it is reasonable to claim that the addition worked as an unnecessary supplement that acted as a gateway for a topic drift described by Warren (2006) as an "almost imperceptible way in which participants in a conversation blur the border between topics as a result of speaking topically rather than speaking on the topic" (p. 166). From line 29 to 30 Trump says that *the war in Iraq was a big fat mistake*, which does not answer the question *do you still believe president bush should have been impeached* on line 8. What Trump does then is he attacks Jeb Bush from line 31 forwards claiming that it took long time for Bush to admit that the war was a mistake. He started hovering over the topic and was able to shift the focus to Jeb Bush by dragging him into the conversation. The addition was not just a way to provide additional information, but a way to perform a topic drift later on.

Trump's answer to the question was clearly non-straightforward, but in addition to that it was nowhere near to being sufficient. He was able to dodge the question by discussing unrelated topics, and when he returned to the actual answer, he was suddenly discussing whether it was a mistake to go to Iraq or not instead of answering did he still believe Bush should have been impeached back then. He began to answer to a question no-one even asked. Further on beyond the excerpt, Judge performs a pursuit and even asks Trump again if he had changed his opinion, to which Trump answers by saying *you do whatever you want, you call it whatever you want*. On top of that the

answer was not pair-conforming, because he did not make a “yes or no”-type answer to the question on any level.

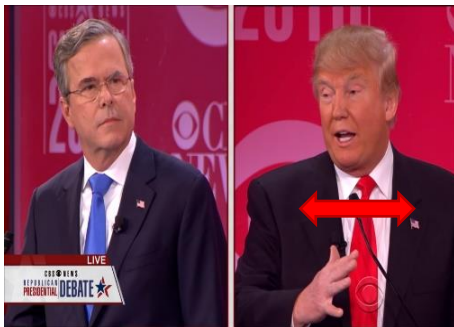


Figure 5a



Figure 5b



Figure 5c

Trump performs a total of three topic shifts and during all those shifts Trump performs a gesture. First one is during the sentence *I have to say as...all over the world* from line 10 to 15. From the figure 5a it can be seen how he moves his right hand in sagittal motion in front of him from left to right twice: first during *with everybody* on line 12 and second during *I've business all over the world* from line 12 to 15. First sagittal sweep refers to *everybody*, multitude of people, and the second one to a wide space *all over*. Second gesture takes place during the *so many people in* on line 18, where he points at the audience with his right hand with an extended arm as can be seen from the figure 5b. The pointing clearly refers to the people in the audience in front of him. Third gesture takes place during the words *so let me just tell you* on line 25 which can be seen in the figure 5c. He raises his right hand, makes an ok sign, breaks the sign and makes a baton gesture. The baton gesture is used as to stress the message to come. Biau & Soto-Faraco (2013) in their research concerning these sort of baton gestures suggested that they have the power unintentionally manipulate the audience to have a more positive and strong reaction to the following message.

All the three gestures take place during topic shifts: first Trump wants to discuss his business, then his funding and finally he returns back to his answer. All the shifts were accompanied by gesticulation, thus, what is common with Trump's answer and the two responses presented earlier in this section is that they all have gestures during the topic shifts.

When looked at the function of the addition it seems that the answer to follow the addition can either be pair-conforming or not pair-conforming and either sufficient or insufficient. The only feature that was present in all the additions was non-straightforwardness, and even when the addition was closely related to the actual answer, as it was in excerpt 3, the initial instinct was that

the addition was not related to the answer. In addition, the topic shifts in the excerpts were all accompanied by gestures that were either related to the rhythm of speech or the addition to come. However, it will be evident later on that the topic shifts seem to be often accompanied by gesture were it an addition, NATQ or CPQ.

#### 4.2. 'Not answering the question'

The second subset discovered is the 'not answering the question' or NATQ. It is defined as follows: a *well/first of all*-prefaced SPP which does not answer to the question presented. The most defining factor here is the possible topic drift or straightforward shift to another topic without intent of returning to the original topic. Otherwise NATQ could be interpreted as participants not sharing the intersubjective field.

The first example of the subset comes from the Democratic debate, where the debaters were discussing whether college tuition should be free or not. Rising costs of college and constant problems with student loan programs had been a major issue in the United States for decades, and naturally it was a part of the discussions during the Presidential debates. Before the debate, Clinton had criticized Sanders' plan of making public colleges free for everyone and she had said, according to Judge, that she is not in favor of making a college free for *Donald Trump's kids*. In the debate Sanders answered a question appointed to him on that issue and after the question Judge turned to Hillary Clinton and started asking her questions. Those questions eventually lead to the discussion about Medicare and Social Security. Consider the following excerpt:

##### (5) *Social security*

1 JUDGE: question was not just about tuition though it was about, (0.3)  
 2 senator's: senator sanders' plan, (0.5) to expand social  
 3 security: to make medicare available to all americans. (0.6) IS  
 4 THAT something you would support? and if not why not?  
 5 (.)  
 6 CLINTON: \*#WELL I- I \*#fully \*fsupport\* social  
 7 cli \*.....\*baton gesture with both hands\*,,,,,,,\*  
 8 fig #Fig. 6a #Fig. 6b

9 CLINTON: security and the most IMPORTANT fight we're gonna have is  
 10 DEFENDING it, (0.3) against continuing republican: [efforts to  
 11 PRIVATIZE it,]  
 12 JUDGE: [DO YOU WANT  
 13 TO EXPAND IT?]  
 14 (0.7)  
 15 CLINTON: I WANT <to: enhance the benefits,> (0.3) for the POOREST  
 16 recipients of social security.

The FPP of the question/answer adjacency pair starting on line 3 is clear and simple enough: *is that something you would support, and if not why?* Judge is asking for Clinton's opinion to senator Sanders' plan to make *Medicare available to all Americans*. After a short pause on line 5, Clinton well-prefaced SPP begins on line 6: *Well, I fully support social security and the most important fight we are going to have is defending it against continuing Republican efforts to privatize it*. It seems as she has dismissed the question and answered to one that was never asked.

One of the key differences between the addition and NATQ is that the participant giving the answer might not perform any sort of expansion at all when they begin their turn as can be seen from the excerpt 5. Clinton was asked a question and she gave an answer right away after the *well/first of all* utterance on line 6 making the answer pair-conforming. However, the question was if Clinton would support Sanders' plan to make Medicare available to all Americans and not if Clinton would support social security. Clinton does not answer the question presented to her. After answering the question briefly from line 6 to 9 she moves on to say from line 9 to 11 that *the most important fight we are going to have is defending it against continuing Republican efforts to privatize it*. She was not asked what the most important fight was. Nevertheless, she decided to comment on that specific issue.

When a debater is not answering a question, a judge may pursue for an answer by making an explicit reference to the original question. In Romaniuk (2013) explicitly referencing the original question was one of the four ways a pursuit could be performed. If Judge pursues for a proper answer, it suggests that the answer was insufficient. During this excerpt Judge does exactly that. From line 12 to 13 Judge interrupted Clinton by asking *do you want to expand it*, because she had already realized that Clinton was not answering the question presented to her. She referenced the original question, because Clinton had either misheard the question or intentionally misunderstood it. Then after a longer pause, Clinton continued to not answer the question straightforwardly by saying that she would *enhance the benefits for the poorest recipients of social security* from line 15 to 16. Pursuits

are a good way of identifying whether the debater had answered the question sufficiently or not, because if they did not, the judge may often continue with a pursuit in order to get the answer out of the debater. Thus, in the excerpt the debater did not answer the question sufficiently.

NATQ begins with the *well*-prefaced response as did the addition. However, in NATQ the response to follow may move right into the answer even though it does not answer the question. Clinton does not perform any transition via expansion between Judge's question and her insufficient answer. Therefore, Judge asks for clarification from line 12 to 13 before Clinton could finish her post-expansion where she started talking about the *most important fight* beginning on line 8. The answer had to be interrupted as from the Clinton was not answering the question presented to her from the point she began answering. After the pursuit, from line 15 to 16 Clinton said she would like to enhance social security for *the poorest recipients of social security* but does not mention directly if she would expand social security. The answer leaves a lot of room for interpretation as she does not directly say yes or no. Beyond the excerpt she goes on to explain that she would help those who are in the most need of social security services, but next debaters' turn explains why Clinton did not address the question. The next debater, Sanders, explained that during the Obama administration the Republicans and some Democrats were in favor of cutting social security until Sanders and some other Democrats created a Social Security caucus to defend it. Clinton's opinion on the topic of cutting social security was not completely clear during the time of the debate and that is the reason why she was not too eager to discuss it. Clearly Clinton did not answer the question during the debate, but afterwards in 2016 she made up her mind and became the advocate for not cutting social security. The background explains well why she gave a non-straightforward and a clearly insufficient answer.



Figure 6a



Figure 6b

While Clinton begins her answer on line 6 by saying *well I*, she prepares both of her hands on for a baton gesture as can be seen from the figure 6a. The gesture peaks at the word *fully* on line 6 as can

be seen from the figure 6b which is followed by retraction and smiling voice during the word *support* on the same line. The baton gesture is used to underline the word *fully*. Biau & Soto-Faraco (2013) in their research concerning these sort of baton gestures suggested that they have are used to emphasize the following utterances to the listeners. Clinton uses the gesture during the word *fully* that is followed by *support social security* which is the point of her argument: she has no interest in abolishing the current social security system.

In the cases of the addition, topic shifts' Gestalt were often composed of gestures and speech especially during the part where the debater used the *well/first of all* utterances. In the excerpts 2 and 3 the debaters used their hands right after saying *well/first of all* during the topic shift to the addition or alternatively during the topic shift back to the original topic or yet another topic as Trump did in excerpt 4 twice. It is evident that during the transitions to different topics the debaters often utilize visuospatial means. For that reason, it is interesting that Clinton used her hands during and after *well/first of all* on line 6.

In the excerpts 2 and 3 the addition that followed the gesticulation did not answer the question. It was only after the addition and yet another topic shift when the debater moved back to answering the question. It is reasonable to argue that Clinton was trying to perform a topic shift to an unrelated topic by not answering the question sufficiently using a pre-second expansion that acts as an answer. Normally a topic shift would be marked by the participant for example by declaring that there is something they would like to say before addressing the question as Sanders in excerpt 3. However, Clinton tried to answer in a manner that would be somewhat related to both the question asked and the further information she wanted to produce after her non-answer.

The answer is therefore an occurrence of topic drift described by Warren (2006) as an "almost imperceptible way in which participants in a conversation blur the border between topics as a result of speaking topically rather than speaking on the topic" (p. 166). Clinton performs the topic drift in her answer and is able to avoid answering the question presented to her. Judge was, therefore, unable to provoke a satisfying response to his question, which lead to her pursuing for an answer during Clinton's answer. Clinton performed NATQ multiple times during the debate and thus two of the best examples of the case are from her turns.

In the following excerpt from the Democratic debate Judge and Clinton are discussing the Russians' involvement in Syria and how the Russians have supported Assad's war against the rebels. The topic

was very prominent during the 2016 elections, because in September 2015, ten months after U.S. joined the war as part of the CJTF-OIR, the Russians had joined Assad's forces in their effort to eradicate the rebellion and the other foreign forces in Syria. Judge is interested if Clinton thinks she had underestimated the Russians when it came to Syria. The fact is that Clinton was the Secretary of State during that time and she might have been able to do something about the issue. First, Judge began talking about Russia in general and moved to his designated question to Clinton. Consider the following excerpt:

(6) *Russians in Syria*

1 JUDGE: you spearheaded the reset with russia did you underestimate the  
 2 russians as- (0.4) and as president, (0.3) what would your  
 3 response to vladimir putin be right now in syria?#  
 4 fig #Fig. 7a  
 5 (0.5)  
 6 CLINTON: +well first of all#, (.) we got a lot of business \*done\* with  
 7 cli +lowers her gaze --> \*briefly looks at Judge\*  
 8 fig #Fig. 7b  
 9 CLINTON: <the eh> RUSSians when Medvedev+ was the president and not putin  
 10 cli raises her gaze back up --->+  
 11 CLINTON: we got a, (0.3) nuclear arms deal, (.) we got the iranian  
 12 sanctions we got, (0.3) an ability to bring important material  
 13 and equipment to our soldiers in afghanistan, (0.3) †there's no  
 14 doubt that‡ when putin came back in and said he was going to be  
 15 president, (0.4) <uh> that did change the relationship. (0.5)  
 16 we have to stand up to his bullying and \*speci<fically in>  
 17 cli \*gazes up and down\*  
 18 CLINTON: syria, \*

From line 1 to 3 the FPP of question/answer adjacency pair was performed when Judge asked Clinton if she thinks she underestimated the Russians in Syria and what Clinton would do concerning the Russian involvement if she got elected. From line 6 to 13 Clinton responded by listing the achievements during her time as the Secretary of State, and then from line 13 to 15 she acknowledged that after Putin was elected for the third time *that did change the relationship*. She never addressed the first question asked by Judge on line 1 which was: *Did you underestimate the Russians?* After saying that Putin's rise to power changed the relationship, Clinton moved on to say



*we have to stand up to his bullying* on line 16 and started discussing what the United States should do with Putin beyond the excerpt. After Clinton had finished her turn, Judge moved to the next question, that was appointed to Bernie Sanders. Judge could have pursued for a more elaborate answer but decided to move on instead.

The *well*-prefaced answer to a question in the excerpt presented is interestingly structured. From line 6 to 13 the answer started with details that were seemingly related to the first part of the question presented by Judge from line 1 to 2: *did you underestimate the Russians?* Clinton acknowledged that Judge was referring to time when she was the Secretary of State from January 2009 to February 2013, because she specifically addressed the detail. Therefore, she must not have misheard the question. May 2012 was the month when Putin became the president for the third time and Clinton began her turn by referring to that. Instead of answering the question if she underestimated Russians, what she did was she gave an answer from line 13 to 15 by saying: *There's no doubt that when Putin came back and said he was going to be president that did change the relationship.* The given answer is not an answer to the question presented. Clinton's answer was insufficient, because saying the relationship would change does not confirm either of the possible sufficient answers of *yes* or *no* either. In addition, the answer is not pair-conforming answer to a yes/no-question, as something what was expected was a yes/no-answer even to some extent. Clinton's answer does not confirm either of the two options.

The excerpt exemplifies that *well*-prefaced answers do not necessarily answer the question sufficiently at all. Furthermore, if there are two questions, one of them may be answered sufficiently and the other may be left unanswered. Clinton answers the second question sufficiently from line 16 forwards and beyond the excerpt. She claims that the current administration is doing great work when engaged in talks with the Russians concerning the situation in Syria. She advocates sending a strong message to the Russians that they have to help the people and would have to be part of the solution instead of *creating more chaos*. However, the first part of the question did not get answered at all. What she did was she brought up the achievements of her time as the Secretary of State from line 5 to 9: nuclear arms deal, Iranian sanctions and ability to give vital support to troops in Afghanistan. There she would have had the opportunity to address the question raised by Judge, but she kept being ambiguous and told Judge that the relationship changed.

As was said earlier, Clinton did not answer the first part of Judge's question sufficiently. From line 5 to 15 she comments on the topic by saying she did not fail when it came to making deals with

Russians and she said that the relationship changed. However, she did not give a straightforward answer the question as from line 16 forwards she started discussing the situation in Syria. Then she began to gaze around instead of focusing her gaze mainly on any specific point right after she said *specifically in Syria* from line 16 to 18. Then she moved on with her answer instead of addressing the question in a straightforward manner. Schegloff and Lerner (2009) discussed the *well*-preface as the first TCU and how it acts as an alert that the following response will be at least somewhat non-straightforward. Clinton's answer to the first question presented by Judge was non-straightforward, however, the answer to the second question was straightforward. She said beyond the excerpt that they have to keep being in talks with Russians and how it was *important to provide safe zones that people are not going to have to be flooding out of Syria*. Therefore, if there are more than one answer following a *well*-preface, it does not always indicate that all the following answers are non-straightforward. It could mean that the first one is, but more data should be gathered.



Figure 7a



Figure 7b

When Judge was asking the question from line 1 to 3, Clinton kept eye-contact with the Judge the whole time as can be seen from the figure 7a. However, after Judge had finished his question, Clinton turned her gaze downwards avoiding eye-contact as can be seen from the figure 7b. She lifts her gaze during the word *done* on line 6, but quickly turns it back down until she said *Medvedev was the president* on line 9. The meaning for avoiding eye-contact is unknown, but it is worth noting that Clinton raised her gaze during the high point of her defense. She was the Secretary of State for four years of which over three were when Medvedev was the president of Russia and one was when Putin was the president. She had under a year time to take part in the communication with Putin. After her defense, she started talking about Syria on line 16. During the words *and specifically in Syria* from line 16 to 18 her speech slowed down and she gazed up and down twice. Interestingly these visuospatial modes take place during the topic shift from Russians to Syria, which is in line with the previous findings in excerpts 2, 3 and 4.

Sometimes, even if the answer is not satisfying enough a judge may still be willing to hear out the whole answer before they begin the pursuit. In the following excerpt from the Republican debate the debaters were discussing amnesty and their take on illegal immigration, both topics always discussed during elections. Cruz had previously promised that he would deport illegal immigrants and reverse Obama's executive action to give temporary amnesty to illegals who were brought to United States by their parents. Judge was interested that if Cruz would be elected, would he use the list of names and addresses created by the temporary amnesty to start deporting illegal immigrants. Consider the following excerpt:

(7) *Door to door*

1 JUDGE: AS PRESIDENT you would have the names and addresses of those  
 2 some \*eight hundred thousand of those that have registered  
 3 *cru* \*nodding occasionally-----  
 4 JUDGE: under that action. hh. (0.4) now you have said that in this  
 5 country we shouldn't go door to door looking for illegals but  
 6 in this case you'd have a list. (0.3) would you# use it?#  
 7 *cru* ----->#  
 8 *fig* #Fig. 8a  
 9 CRUZ: \*(1.8) well# you know your question\* highlights a- a sharp  
 10 *cru* \*gazes downwards towards the podium\*  
 11 *fig* #Fig. 8b  
 12 CRUZ: difference on immigration on this stage. (0.5) you know in a,  
 13 (.) republican primary everyone talks tough, (0.3) on  
 14 immigration. (.) everyone is against illegal immigration in a  
 15 republican primary. (0.7) BUT AS VOTERS we've been burned. (0.3)  
 16 over and over again by people that give us a great campaign  
 17 °speech and they don't walk the walk°.  
 18 ((lines omitted))  
 19 CRUZ: and led the fight, (.) to defeat THAT amnesty plan.  
 20 (2.4)  
 21 JUDGE: so would you, (0.4) would you use, (0.5) the [addresses?]  
 22 CRUZ: [now that] moment,

From line 1 to 6 Judge is setting up his upcoming question on line 6 *would you use it* referring to the list of illegal immigrants which Cruz would get his hands on if he got elected. After the FPP of the question/answer adjacency pair was performed, Cruz began his turn on line 9. He did not start by

giving an answer to the question presented which would have been the SPP of the adjacency pair. Instead, he claimed from line 9 to 12 that the *question highlights a sharp difference on immigration on this stage* and started discussing how all candidates are against immigration during the primaries but have always failed to deliver if elected from line 12 to 17. He continues to discuss this topic, but never answers Judge's question. Finally, after Cruz had finished his turn on line 19, Judge asked Cruz again *would you use the addresses* on line 21.

While in excerpt 6 Clinton did not perform a pre-expansion for her new topic, in Cruz's case there is a pre-expansion but no answer of any sort. Question *would you use it* has two possible straightforward and pair-conforming answers to it which are *yes* or *no* followed by explanations. However, on line 9 Cruz goes straight into the pre-expansion *well, you know your question highlights a sharp difference on immigration on this stage* that prepares for his future turn beginning on line 12. He dismisses the opportunity to answer the question and moves to his turn about immigration and primaries. Cruz's turn is indisputably non-straightforward and in addition not pair-conforming, because during his turn he does not address the question at all. Cruz does not state if he would use the list of addresses or not and decides to talk about the differences between the candidates instead. He was more interested in telling how candidates tend to speak tough on immigration during the primaries, but afterwards would not deliver. Schegloff and Lerner (2009) suggested that the answers following a *well*-preface as a TCU would be to some extent non-straightforward and the preface would act as an alert of this non-straightforwardness. However, if the participant does not address the question at all, as Cruz in this excerpt, calling it an alert in every case is unsatisfactory, because alerting someone of "I am not about to answer your question" is illogical. Therefore, *well*-preface should be referred to as an indicator of non-straightforwardness that does not rely on the participants' intent.

The fact that the question was not addressed suggests that the answer is insufficient. In addition, Judge's verbatim repeat pursuit, which was one of the forms of pursuit listed by Romaniuk (2013), on line 21 *would you use the addresses* discloses the case that the answer was insufficient. Judge asked the question again, because it was not answered. This excerpt has all the properties of a failed answer: non-straightforward, not pair-conforming, insufficient and followed by a pursuit.

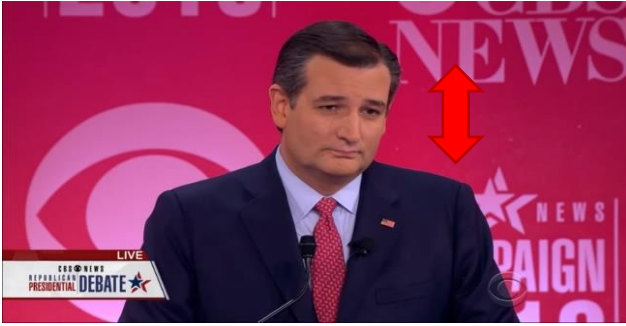


Figure 8a

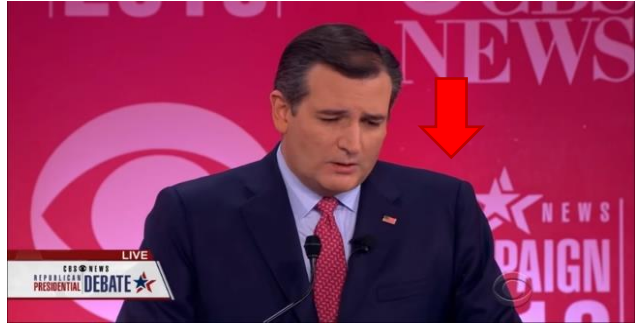


Figure 8b

In excerpt 6 Clinton lowered her gaze while she was beginning her turn. In this excerpt, Cruz keeps nodding his head up and down during Judge's turn as can be seen from the figure 8a. When Judge finishes her turn on line 6, Cruz instantly lowers his gaze as in the figure 8b, stays silent for 1.8 seconds, starts talking and raises his gaze after the word *question* on line 9. Clinton used her gaze during a topic shift in excerpt 6 and so does Cruz in this excerpt, too. He skips answering the question in a straightforward manner and moves to his pre-second expansion which is first accompanied by shifting the gaze down and back up. Visuospatial means are again used during a topic shift.

As was noted in 4.1. *Addition* the *well*-prefaced answers are often accompanied by gestures during the first utterances. In NATQ the debaters began their *well*-prefaced answers quite often with lowering their gaze (not limited to excerpts 5 and 6), which is intriguing. However, going into more detail on the subject should be a topic of another research, but it is worth noting that the beginning of the *well*-prefaced answers of NATQ are also accompanied by other visuospatial modes than gestures. It must be declared that it is not suggested in this study that lowering gaze is related to lying or deception. According to Sporer (2014) there is no connection between deception and avoiding eye-contact or decreasing and increasing the amount of any other visuospatial mode whatsoever. It is a common misbelief which has been debunked by multiple meta-analyses on the subject. What is suggested is that there might still be a connection which is not related to lying but to turn preparation instead.

NATQ turns either begin with a non-answer as in excerpt 5 or with an expansion as in excerpt 6. All the answers were logically non-straightforward, because if one is not about to answer a question, the answer will never be straightforward. All the answers were insufficient, because either Judge would follow the answer with a pursuit or elaborating question or a debater was somehow able to dismiss the question entirely. However, excerpt 4 exemplified that the turn can be pair-conforming even if it would not answer the question. A debater could answer a question, but the answer would

not be related to the question presented. This was an interesting tactic to answer as it worked as a great way to move to another, safer topic.

#### 4.3. 'Commenting on a prior question'

The third subset 'commenting on a prior question', hence CPQ, is defined as follows: when a judge asks debater a question, instead of giving a sufficient answer, they comment on or answer to a prior question presented to a debater prior to the current question. This style of commenting suggests that a debater had a comment or answer on their mind which they had no opportunity to say out loud during the appropriate time, because they never got the opportunity. For a debater the only way to get the opportunity to comment on a prior question is to somehow shift the topic to the prior question. The first excerpt exemplifies the most common way to do this.

During the Republican primaries in 2016 one of the key questions was how the candidates would deal with ISIS if they were to be elected the president of United States. One worrying detail to many republicans Ted Cruz's reluctance towards the idea of putting American boots on the ground when fighting ISIS. Instead, he was on the side of the idea of arming the Kurds to fight for them in the northern areas of Syria. In the excerpt to follow, Judge asks a question considering Cruz's prior statements. Consider the following excerpt:

##### (8) *Ground troops*

1 JUDGE: JUDGE: you talked about the first >the first gulf< war as  
 2 being a kinda model for your focused and determined effort to  
 3 go after a- after isis, .hh (0.5) but ↑there were seven hundred  
 4 thousand ground troops as a part of that↓ so and you don't have  
 5 a <ground component to your plan> why?  
 6 (0.9)  
 7 CRUZ: #well we need to focus on what \*the# \*objective\* is\*.  
 8 cru \*.....\*baton gesture\*,\*,\*  
 9 fig #Fig. 9a #Fig. 9b  
 10 CRUZ: (.) \*+>you know#< your+ QUEStion\* about the first three questions  
 11 cru \*moves his left hand back and forth in a sagittal motion\*  
 12 +gazes upon his lower left side+

13 fig #Fig. 9c

14 CRUZ: you would ask in in the, (0.7) situation room, .hh (0.6) I think

15 it is a problem, (1.0) if the president commander-in-chief we

16 we've elected does not have the experience and background to

17 understand the threats facing this country coming in on day

18 one. (0.5) if you look at the threats facing this country the

19 SINGLE gravest threat, (0.9) NATIONAL security threat, (0.3) is

20 the threat of a nuclear iran that's why i pledged on day one,

21 .hh (0.3) to rip to shreds this iranian nuclear deal and anyone

22 that thinks, (.) you can NEGOTIATE with KHameinei, (0.8) does

23 not understand the nature of khameinei. (0.5) \*when\* it comes

24 cru \*baton gesture with his right hand\*

25 CRUZ: to isis,

26 (lines omitted)

27 JUDGE: very quickly, (0.5) thirty second follow-up you- you've said

28 you would- >that the essentially the< kurds would be the

29 American ground forces in there..hh (.) the- the- the criticism

30 that the experts have on that is that the kurds, (0.4) only can

31 work within their territory if they take larger amounts of

32 territory, .hh (.) you have an ethnic war with the arabs so the

33 ku- the kurds cant really, (.) do as much as you seem to be

34 putting on, (.) on their backs.

From line 1 to 4 Judge brings up the fact that Cruz had said in the past that the Gulf War was a kind of a model for his future plans when going after ISIS. During the Gulf War there were 700 000 ground troops as part of that plan, when Cruz's plan against the ISIS has no U.S. ground component in it, and as the FPP Judge asks Cruz *why* on line 5. After Cruz briefly gives an answer as the SPP that *well, we need to focus on what the objective is* on line 7, he started talking about implications of Trump's answer to a prior question and what was wrong with Obama and Clinton's plans against the ISIS from line 10 to 23. After that, Cruz told Judge that as a substitute for not having U.S. ground troops he would arm the Kurds, and he sees that sending U.S. ground troops should be used as the last resort appointed by army officials. That was the claim Judge wanted on the record as he does a follow-up from line 29 to 33 noting: *The criticism the experts have on that is that the Kurds only can work within their territory. If they take larger amounts of territory, you have an ethnic war with the Arabs, so, the Kurds can't do as much as you seem to be putting on their backs.*

For the analysis we will move back to the beginning of Cruz's answer. It is obvious that Cruz knows he shifted away from the original question, because he performs two separate expansions, one before and one after his comment to the prior question: the pre-second insert expansion on line 10 *you know* and another pre-second insert expansion from line 23 to 25 *when it comes to Isis*. He had to perform expansions in order to follow the rules of conversation, which suggests that he was aware of himself changing the topic. The first expansion after the brief answer to the question on line 7 marks the shift in discussion as Cruz wants to comment on something before giving a more elaborative answer than just: *well, we need to focus on what the objective is*. Cruz's pre-expansion *you know your question* sounds if something just hit him and he wanted to elaborate on that. Cruz's actions cause a topic shift to occur, thus, turning the focus of the discussion on one debater's inadequateness to be a president and his inability to recognize the national threats facing the country. The comments are related to other debater's answer to a question, to which Cruz had an appropriate comment for. Cruz found the opportunity to add his comment to the conversation by briefly addressing the question presented by Judge and then utilizing an expansion to return to the other topic. After Cruz had had his say, the second expansion *when it comes to ISIS* from line 23 to 25 is there to mark a topic shift back to the original topic of his plans against ISIS. Cruz's second expansion suggests that it was necessary to return to the topic and a more elaborate answer to the original question was about to occur.



Figure 9a

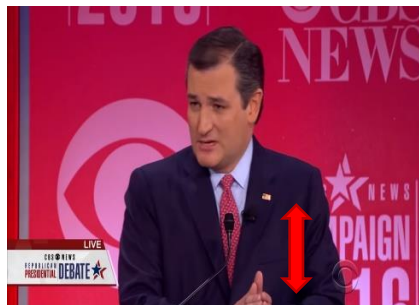


Figure 9b

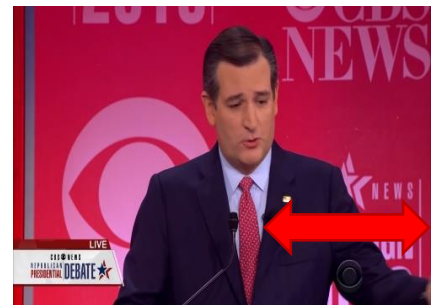


Figure 9c

When Cruz addressed Judge's question by answering the initial question briefly on line 7, the comment was accompanied by a baton gesture, which apex was during the second syllable of the word *objective*. The baton gesture took place in front of Cruz right above the podium. He moved both of his arms into a prayer like position as can be seen from the figure 9a and the baton gesture can be seen from the figure 9b. After the short answer, Cruz began to answer to a prior question accompanied by another phase of the gesture as can be seen from the figure 9c. The sagittal back



and forth motion performed by his left hand, at the time he switched the focus to a prior question while he said *you know, your question* on line 10, looked similar to moving something from outside the podium in front of the participant. This same location of focus was referenced during the baton part of the gesture, which is interesting as it seems like the old topic was in some sense “brushed off” the podium by a new one. The fact is supported by the participant’s speech as he introduces the prior topic back to the discussion after giving his initial answer. Cruz’s second expansion on line from line 23 to 25 was accompanied by a gesture, too. During the word *when* he makes a baton gesture with his right hand as an emphaser of his speech. The case of CPQ is not different from the addition as they are often surrounded by expansions, too. In addition, their Gestalt was composed of speech and gestures, as well. This will be further discussed when analyzing the next excerpt.

As Schegloff and Lerner’s (2009) study suggested, *well*-prefaced answers to *wh*-questions that are pair-conforming are not necessarily sufficient answers. Cruz’s gave a pair-conforming answer on line 7 *well, we need to focus on what the objective is*. The answer follows the conventions of question/answer adjacency pair, but the answer itself was insufficient. Even Cruz’s own action proves that, because he himself elaborated his answer further from the line 23 forward. However, the elaboration was not enough, because from line 27 to 34 Judge asked a follow-up concerning Cruz’s answer.

The truly interesting question is what made Cruz choose to comment on a prior topic first from line 8 forward instead of elaborating his insufficient answer first. This is typical for the CPQ category as will be seen from the next excerpt, too. Schegloff and Lerner’s (2009) study also suggested that *well*-prefaced answers sometimes suggest that the answer to follow is not always straightforward. In CPQ most of the time the question is addressed in a straightforward manner, but the initial answers tend to be quite short and are often followed by further elaboration after the CPQ. It is fair to say that the answers with a CPQ are not as straightforward as they could be, and therefore, they are classified as non-straightforward, even though the question is addressed and elaborated before a judge is able to ask further questions.

Cruz example portrays the usual form for CPQ that has three phases: briefly answering to the question, CPQ and finally elaborating on the answer. Another example of the form of CPQ can be seen in the next excerpt from the Republican debate, as well. Before the excerpt Judge raised up the fact that Carson had said during a Fox debate that he had had more 2 am phone calls than any other candidate on the stage. Carson had said that he had had to make life and death decision and

had put together teams to make something unheard of to happen before. However, Judge said that it did not matter in the context of political decision making, because during those phone calls Carson was dealing with situations related to his expertise as a physician, which he had decades of experience with. Consider the following:

(9) *2am phone calls*

1 JUDGE: when the two am phone call comes. <so isn't that a liability?  
 2 (0.5) ((Carson smacks his lips)) (0.3)  
 3 CARSON: UH no it isn't. (.) uh first# of \*all\* let me, (.) uh, (0.3) go  
 4 car \*raises his right-hand index finger upwards\*  
 5 fig #Fig. 10a  
 6 CARSON: back to your, (.) uh first question for me. (0.5) uh it <wasn't  
 7 phrased as who gets to,> (0.3) u- nominate supreme court, (0.6)  
 8 uh appointees of course that's the president so, .hh (0.3) i  
 9 know that there are some left-wing media who'd try to make hay  
 10 on that. .hh (0.5) UHH \*SECONDLY\*#, (0.4) UH, (.) THANK you for  
 11 car \*shakes hands once in back and forth motion\*  
 12 fig #Fig. 10b  
 13 CARSON: including me into the debate. (.) two questions already this is  
 14 great. .hh (0.5) ((Carson laughs for a 1.1 seconds)) (0.3) now  
 15 as FAR,  
 16 AUDIENCE: ((Cheering for 2.3 seconds))  
 17 CARSON: ((Carson smacks his lips)) (0.7) \*#as far as\* those two am phone  
 18 car \*spreads his hands in sagittal motion outwards\*  
 19 fig #Fig. 10c  
 20 CARSON: calls are concerned,

Before the excerpt Judge underlays his question by contesting the idea that Carson's experience as a physician would be comparable to one's experience as a politician when concerning 2am phone calls. The question to which Judge expects an answer is asked on line 1: *So, isn't that a liability?* Carson answers the question briefly on line 3: *No, it isn't.* Then he performs a post-second insert expansion starting on line 3 where he marks the topic shift to the prior question asked by Judge prior in the debate: *First of all, let me go back to your first question for me.* He takes the conversation way back into the debate, because other participants had commented on his answer to a prior question and he decided to comment on the comments. It is not a rebuttal that follows the rules of the debate, because Judge did not specifically ask for Carson's rebuttal. Carson then reminds the

listener of what the question was exactly and corrects his own prior statement. Next on line 10 he says *secondly* to mark a topic shift and from line 10 to 14 he comments on him not getting enough attention in the debate and starts laughing. Finally, in the end from line 15 to 20 he marks yet another topic shift with a third expansion *now as far as, as far as those 2am phone calls are concerned* that prepares the listener for his elaboration for the initial answer *no, it isn't* on line 3. The form of CPQ in this excerpt is similar to the form of the CPQ in excerpt 8.

After answering the question briefly on line 3, Carson marks the topic shift with the phrase *for first of all* instead of beginning his whole answer with it. When compared to the excerpt 8 discussed earlier, the similarity is that Carson answers the question right away pair-conformingly before performing the *well /first of all* expansion that would lead to a comment to a prior question. The form is the same in this excerpt. However, after the CPQ Carson performed yet another topic shift which lead to an addition about him not getting enough time in the debate. In excerpt 4 Trump performed a string of additions which were not related to each other. In this excerpt the first unrelated topic was a CPQ which was followed by a case of the addition. The two can be utilized in the same manner.

In this case Carson's initial answer was undoubtedly insufficient even though it was yet again pair-conforming. After the second topic shift beyond the excerpt, the candidate answered the question sufficiently. He specified that it is his judgement and experience that allows him to make good decisions when faced with new problems that require brand new solutions that is his cutting edge when it comes to those situations. Judge did not ask any further questions which supports the analysis.

The most prominent difference between the excerpt 8 and 9 is that after the debater has answered the question Judge either asked or did not ask a follow-up question from the debater. Thus, answering the question insufficiently at first and moving to comment a prior question does not determine if the elaboration of the insufficient answer will be sufficient itself. Whether Judge will or will not ask a follow-up is not dependent on a debater's utterances. Rather on Judge's read on debater's answer.



Figure 10a



Figure 10b



Figure 10c

This excerpt is a good example of how topic shifts can be plenty and accompanied by gestures. Carson's Gestalts during the three shifts between topics, during the phrases *first of all* on line 3, the phrase *secondly* on line 10 and the phrase *as far as* on line 17, were all composed of gesticulation and speech. During the words *first of all* on line 3 Carson raises his index finger as seen from the figure 10a, during the word *secondly* on line 10 he moves his hands in a back and forth motion as in the figure 10b, and during the words *as far as* on line 17 he moves his hands in sagittal motion as seen in the figure 10c. Carson's topic shifts are accompanied by gestures as was Cruz's shifts in excerpt 8, which suggests that topic shifts in CPQ are often accompanied by gestures, too. In addition, gestures in CPQ are not limited to topic shifts that are *well/first of all*-prefaced. It seems as any topic shift during answers are accompanied by gesticulation.

The last excerpt of this category exemplifies how the participant can comment on a prior question and at the same time avoid answering the original question by not performing a focus shift back to the original question. The excerpt is from the Democratic debate. Before the excerpt Judge and Sanders had been discussing free public colleges and how Sanders' plan would substantially lower interest rates of college debts. After Sanders had finished his turn, Judge turned to Clinton and decided to ask her take on Sanders' plan to expand social security and giving all Americans Medicare. Consider the following excerpt:

#### (10) College tuition

- 1 JUDGE: and SECRETARY, (0.3) SECRETARY clinton it's NOT just college  
 2 tuition. (0.4) that senator sanders is:, (0.3) talking about  
 3 expanding social security, (0.3) and giving all americans  
 4 medicare. (0.3) what's wrong with that?  
 5 (0.3)  
 6 CLINTON: \*well#\* \*LET\* \*me address:\* uh college affordability

7 cli           \*.....\*raises both hands\*,,,,,,,,,,,,,,\*  
 8 fig                 #Fig. 11a  
 9 CLINTON:   because I HAVE A- a- (0.3) a plan that i think will really,  
 10                 (0.5) uh zero in on what the problems are first, (0.8) \*#ALL\*  
 11 cli   \*brushes her right hand\*  
 12 fig   #Fig. 11b  
 13 CLINTON:   the 40 million americans who currently have student debt will  
 14                 be able to refinance their debt to a low interest rate.

On line 1 Judge is pulling Clinton into the discussion by starting with the words *and secretary*, *secretary Clinton* of which especially the word *and* is suggesting that her question will be related to the discussion she had with Sanders few seconds ago. From line 2 to 4 she lays out the background to her question *senator Sanders talking about expanding social security and giving all Americans Medicare* and presents the question *What's wrong with that?* The FPP being a question requires an answer as the SPP. Clinton begins her *well*-prefaced answer on line 6 with the words *well, let me address college affordability* that was not the topic Judge was asking about. From line 9 to 10 Clinton presents what she is going to discuss and from line 10 forwards she discusses how her plan is going to *zero in on* the problems of college debt. Beyond the excerpt she keeps talking about the college debt for the rest of her turn and does not return to the original topic.

After listening to Judge's question, Clinton moves right away to the topic shift on line 6. She marks the topic shift with the pre-expansion *well, let me address college affordability*. She does not take a position to either social security or Medicare, but instead decides to move straight to the prior question about college debt. She does not begin her answer pair-conformingly, which would have been to answer the question presented. She wanted to discuss another topic and in order to do it she had to perform the topic shift. Thus, she had to mark the topic shift right away, which is not a pair-conforming way of starting the turn following a question. Pair-conforming would have been to answer the question presented to her. Rest of the factors are easy to determine, because Clinton did not answer the question at all during her turn and in addition Judge pursued for the answer after Clinton's turn. Therefore, her answer was insufficient and of course non-straightforward, because she did not go immediately to the answer. Similar to Cruz's use of the *well*-preface in the excerpt 7, Schegloff and Lerner's (2009) claim of the preface alerting the questioner of the following response being non-straightforward does not apply in this excerpt either. Hillary alerting that she is not going

to respond the question in a straightforward manner does not make sense here either. It should be called an indicator of non-straightforwardness.

The key difference between this excerpt when compared to excerpts 8 and 9 is that the debater did not return to the question after the CPQ. In fact, after the topic shift and beyond the excerpt one can hear that the Judge is talking over and trying to pursue for the answer, but the debater keeps going with her answer. Therefore, when it comes to CPQ it is in the hands of the debater whether they want to answer the original question or not. The CPQ can act as either a momentarily detour that will return to the topic at hand as in excerpts 8 and 9 or CPQ can act as a way to spend the given time on a question which was not even ever presented.



Figure 11a



Figure 11b

In excerpts 8 and 9 when the CPQ's topic shifts were preceded by short answers they were accompanied by gestures. In this excerpt, even though not preceded by a short answer, topic shifts were accompanied by gestures, too. After Judge had asked her question on line 4, in the beginning of the topic shift on line 6 Clinton gesticulates with her hands in a sagittal up/down motion as seen from the figure 11a. The gesture is prepared during the word *well*, apexes during the word *let* and retracts during the words *me address* on line 6. The information is in line with the findings of the excerpts 8 and 9. However, a more interesting detail is the part where Clinton starts explaining the details of her plan. On line 10 Clinton gesticulates by brushing her hand in a lateral motion during the word *all* on line 10 before moving to the more specific explanation of how her plan would affect the student debt issue as seen from the figure 11b. Changing on what level of detail one is speaking cannot be interpreted a topic shift but there is a shift from general to details. There is a shift from a more general way of speaking of the measures to the details of the measures, and Clinton does not gesticulate during the shift but after it. Therefore, it is interesting to notice that just any shift is not enough for gesticulation during the phrase that initiates a shift, but it seems as it has to be a topic shift specifically. However, further research should be conducted on this topic.

One might argue that CPQ is so similar to the addition and that separating them to two different subsets would be unnecessary. However, there was one major dividing factor between the two. The difference between CPQ and addition was that CPQ never enhanced the answer while addition on the other hand did. Even if the debater shifted the topic after CPQ, the CPQ was never related to the answer to follow. The answers to the questions would not be affected if the CPQ was not present during the debaters turn.

## 5. Discussion and conclusion

The reason why political debates are held is to give the voters a picture of the ideals and policies the candidates advocate for. It is in the debaters' interest to paint a picture of themselves that is compelling to the maximum amount of audience and brings more voters in than more voters out. Audience is most of the time the passive actor in the debate setting. Even though the audience is responsible for giving the votes after the debates and they may even occasionally intervene with the debates, they rarely contribute to them. It is the judges' work to act on behalf of the audience. They are the other active participant of the debates even though they are not the ones the information is appointed to. While debaters are trying to outshine their opponents, the judges' task is to rouse discussion that would reveal the true intentions and flaws of the debaters. Judges are usually journalists as they have every incentive to expose the debaters' true intentions. Therefore, when debaters are not answering a question straightforwardly the judges should carefully listen to what the debaters are saying. Otherwise they are in the risk of losing track of what the debaters are saying which could enable opportunities for the debaters to not answer the questions presented to them. Because the results of this study revealed new information about debaters answering non-straightforwardly, the results could be utilized by judges in future debates.

In this study, it was revealed that when the debaters did not answer a question straightforwardly to the best of their abilities, they either added a comment not related to the question, avoided answering the question or commented on a prior question. The first subset called 'addition' had three different uses in a debate. First one was a circumbendibus to the original topic as in excerpt 2, where Carson wanted to comment on the word *politician* Judge used. Carson told the audience that he is never going to become a politician. Carson repeated the statement he had made before as to differentiate himself from the other debaters, excluding Trump, who were seen as more traditional politicians than himself. After making the statement Carson shifted back to the original topic. Second use for the addition was seen in the excerpt 3 where Sanders was able to pull another debater into the turn he had via the addition. He applauded a co-debater Webb for his achievements in the military. Then Sanders shifted back to the original topic and answered Judge's question revealing that he had worked with Webb on issues concerning the education and health care of military veterans. Sanders was able to give additional information that benefitted his upcoming



argument about him being an adequate candidate to becoming the commander-in-chief. Third type of use was seen in excerpt 4 where Trump used multiple additions in order to start discussing something completely else than what was the subject before the additions. He used the additions about him as a man who gets along with everybody, who knows a lot of people and who is a self-funder as kind of a diversions. He did not go back to the original question of did he think George Bush should have been impeached or not. The additions led to Trump discussing what a mistake the Iraq War was and how Bush did not condemn the war early enough. However, the diversion was detected by Judge who pursued the answer after Trump's turn.

It might initially seem like not answering a question is a black-and-white subject, but nothing could be further from the truth. The second subset called NATQ illustrated that participants not answering a question is more like a scale with two extremities: debaters either speak topically on the subject or do not answer the question at all, and then move to different topic. In the excerpt 5 Judge presented two questions to Clinton. She was able to discuss the first question topically and then moved on to the second question without answering the question at all. In the excerpt 6 Clinton gave a brief answer that did not answer the question at all and quickly moved on to a subject more of her liking. Cruz on the other hand did not even address the question before moving to a different subject in the excerpt 7. All the answers were non-straightforward and insufficient, however, the differentiating factor between the NATQs was how pair-conforming the answers were. It depended on whether the debater answered a question or not. It is a matter of intersubjective field and whether it can be indisputably proven that the debater understood the question correctly. Here is a clarifying graph:

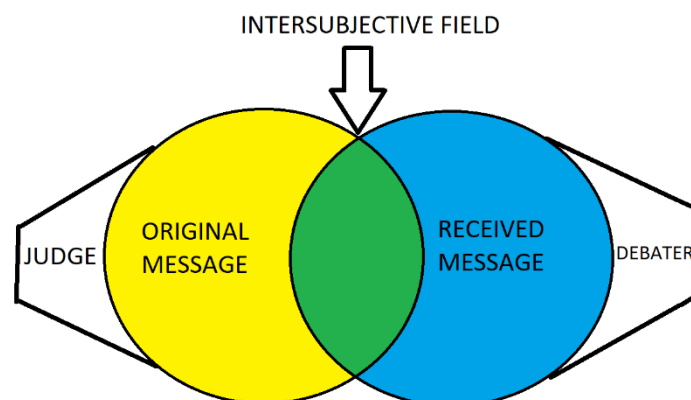


Figure 12

Figure 12 is here to portray how the question can be misunderstood by the debater. After Judge has presented their question, the message is separated from the idea. The original message enters the intersubjective field from which it cannot reach the debater without them interpreting it. Judge has no longer control over the message as it is now debater's task to interpret it. After interpreting the message, the debater receives it and answers as they see fitting. What cannot be reveal by using CA without interviewing the participants is if the original message and the received message are the same. Therefore, whether Clinton misunderstood the question or intentionally did not answer it in excerpt 6 is beyond the reach of this study. This is why an answer may be wrong and at the same time pair-conforming. Someone may answer a wrong question without acknowledging it. Therefore, the focus should be on the turn and not on the utterances when observing whether an answer is pair-conforming or not.

The third subset called CPQ portrayed that the participants might want to comment on a question or a subject, but because they are not given the opportunity to do so during the time, they have comment on it on their own turn. Most often the debater briefly addressed the question presented to them as in excerpts 8 and 9. However, that generalization would be incorrect, because sometimes the debater would not even address the question presented before shifting topics as Clinton did in excerpt 10 which caused the answer to be insufficient, however, the initial answers in excerpts 8 and 9 were insufficient, too. The intention of the debaters showed that even if they would answer the question briefly, they still would elaborate on it later. A conclusion can be drawn that an answer before CPQ is most often insufficient and will more often be elaborated or pursued later than not.

Of the three subsets the addition had the widest of variety functions. A more precise analysis on the addition should be conducted as there was too much variations between the instances. Dividing the addition into more subsets would help defining what the ways are to give additional information before answering a question, because now the answer is that there are multiple options.

NATQ and especially CPQ did not have as vast variety as the addition. In NATQ the debaters never gave a sufficient answer and in CPQ the debaters almost always returned to the original question, the exception being Clinton in excerpt 10. For future research both of the two categories should be treated separately and NATQ could be researched from the perspective of what are the successful ways to perform the NATQ that did not evoke a pursuit and what were the unsuccessful ones. The findings of such study could have real applications in journalism or political debates overall.

A visuospatial finding that was discovered was gesticulation during topic shifts. Whether it was a case of the addition, NATQ or CPQ, the debaters often gesticulated during topic shifts to new topics or when they performed topic shifts that lead to the original topic. They were used to either parse speech, reference to surroundings, perform abstract tasks or to emphasize an upcoming point. The latter was the most prominent of the bunch, and the finding seems to be in line with Krahmer & Swerts (2007) findings on visual beats (nod, hand movement, raising eyebrows) just before uttering a word. They found that if visual beat was performed right before a word, the perceived prominence of the word increased in the listener's opinion. It is reasonable to say that if one desires to put more focus on another subject, utilizing visual beats would be a great idea. The intertwining of gestures and topic shifts came up in the analysis so often that it should be a topic of its own in future research. In this study the surface of the topic was merely scratched, and a collection of vast volume should be gathered.

Another interesting visuospatial mean utilized in the excerpt was gaze during the subset NATQ. Clinton in excerpt 6 and Cruz in excerpt 7 turned their gaze downwards right after they were asked a question they did not answer properly. It was not related to a topic shift, because even though Cruz's gaze turned downwards during a topic shift, Clinton's gaze turned downwards when she was answering the question. Claiming that turning the gaze downwards is connected to debaters not answering the question is not a reasonable assumption to make with this amount of data. However, further studies on the topic could be conducted that would find if there is a connection between the two or not. Knowing if a debater is more likely to not answer a question after turning gaze downwards could be information worth pursuing.

Interestingly, nothing in any of the CPQ excerpts suggested that the participants tried to buy time for their answer by commenting something else first. They either answered the question afterwards or they did not even answer the question as Clinton in the excerpt 10. It seemed as they just had the comment prepared in their mind and regardless of the question and they decided to comment first before elaborating. However, in the case of the addition the irrelevant commenting seemed enable postponing the answer to the question or even lead to avoiding the question in its entirety. The postponing was visible in the excerpt 2 as multiple repairs after the addition either suggested that Carson was rephrasing his answer, or he was trying to think of an answer to the question. Avoiding the question in its entirety was very present in the excerpt 4 where Trump kept rambling on how he knows people and how he gets along with anybody, after which he started accusing Jeb

Bush. It kept going on till the judge asked his question again to which Trump did not answer. However, the addition is not an indicator of purely postponing or not answering, because in excerpt 3 Sanders' addition was related to his later answer. He was setting up his answer and not trying to slow down the debate. Therefore, the addition should be further researched as it seems as there are too many different functions for the addition with the definitions of this study. To conclude anything definitive of this subset would go too far into hypothesizing for a conversation analytic study.

This study is not broad enough to draw any final conclusions on this topic, because the subsets were too wide, and a more precise study should be conducted. However, the goal of this study was to reveal some aspects of political debates that have not yet been researched. That goal was achieved as *well*-prefaced non-straightforward answers were discussed in as large scope as the size of this study enables. The three subsets revealed multiple different uses for *well*-prefaced expansions and an interesting aspect of the visuospatial means during topic shifts was discussed, too. Multiple possible options for future research were also suggested, thus, a good basis for conversation analytic studies of political debates was laid.

## References

- Andone, C. (2013). *Argumentation in political interviews: Analyzing and evaluating responses to accusations of inconsistency*. Amsterdam: John Benjamins Publishing Company.
- Atkinson, M. (1984). *Our masters' voices: The language and body language of politics*. London: Methuen.
- Ballard, E. G. (1971). *Philosophy at the crossroads*. Baton Rouge: LSU Press.
- Biau, E., & Soto-Faraco, S. (2013). Beat gestures modulate auditory integration in speech perception. *Brain and Language*, 124(2), 143–152.
- Braybrooke, D. & Rice, K. D. (2006) *Analytical political philosophy: From discourse, edification*. Toronto: University of Toronto Press.
- Chilton, P. (2004). *Analysing political discourse: Theory and practice*. New York: Routledge.
- Cohen, H., Comrie, B., & Lefebvre, C. (2013). *New perspectives on the origins of language*. Amsterdam: John Benjamins Publishing Company.
- Drew, P., & Heritage, J. (1992). Analyzing talk at work: An introduction. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 3–65). Cambridge: Cambridge University Press.
- Emmertsen, S. (2007). Interviewers' challenging questions in British debate interviews. *Journal of Pragmatics*, 39(3), 570-591.
- Enfield, N. J. & Sidnell, J. (2017). *The concept of action*. Cambridge: Cambridge University Press.
- Essed, P., Farquharson, K., Pillay, K. & White, E. J. (2018). *Relating worlds of racism: Dehumanisation, belonging, and the normativity of European whiteness*. London: Palgrave Macmillan.
- Fitch, K. L., & Sanders, R. E. (2005). *Handbook of language and social interaction*. Mahwah: Psychology Press.
- Goffman, E. (1963). *Behavior in public places: Notes on the social organization of gathering*. New York: Free Press.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Greatbatch, D. (1986). Some standard uses of supplementary questions in news interviews. In J. Wilson & B. Crow (Eds.), *Belfast working papers in language and linguistics* (Vol. 8, pp. 86–123). Jordanstown: University of Ulster.
- Have, P. t. (2007). *Doing conversation analysis*. London: Sage Publications Ltd.

- Heritage, J. & Clayman S. (2010). *Talk in action: Interactions, identities, and institutions*. Chichester: Wiley-Blackwell.
- Hutchby, I. & Allan, S. (2006). *Media talk: Conversation analysis and the study of broadcasting*. Maidenhead: McGraw-Hill Education.
- Jarmolowicz-Nowikow E. (2014) Pointing by hand: Types of reference and their influence on gestural form. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill and J. Bressemer (Eds.), *Body – Language – Communication: Volume 2 : An international handbook on multimodality in human interaction* (pp. 1824-1831). Berlin/Boston: Walter de Gruyter GmbH.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13-31). Amsterdam: John Benjamins Publishing Company.
- Jeffrey D. Robinson (2014) What “what?” tells us about how conversationalists manage intersubjectivity. *Research on Language and Social Interaction*, 47(2), 109-129.
- Jucker, A. H. (1986). *News interviews: A pragma-linguistic analysis*. Amsterdam: John Benjamins Publishing Company.
- Kendon, A. (1990) *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Kitzinger C. (2013.) Repair. In Sidnell J. & Stivers, T. (Eds.). *The handbook of conversation analysis* (pp. 229-256). Chichester: Wiley-Blackwell.
- Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3), 396–414.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Ladewig, Silva H. (2014) Recurrent gestures. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill and J. Bressemer (Eds.), *Body – Language – Communication: Volume 2 : An international handbook on multimodality in human interaction* (pp. 1558-1574). Berlin/Boston: Walter de Gruyter GmbH.
- Lausberg, H. (2014) Body movements and mental illness: Alterations of movement behavior associated with eating disorders, schizophrenia, and depression. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill and J. Bressemer (Eds.), *Body – Language – Communication: Volume 2 : An international handbook on multimodality in human interaction* (pp. 1905-1913). Berlin/Boston: Walter de Gruyter GmbH.

- Mondada, L. (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, 9(2), 195–226.
- Mondada, L. (2014a). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65, 137-156.
- Mondada, L. (2014b). Cooking instructions and the shaping of things in the kitchen. In T. Heinemann, P. Haddington, M. Nevile & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 199-226). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20, 336–366.
- Moore, R. J. (2015). Automated transcription and conversation analysis. *Research on Language and Social Interaction*, 48, 253–270.
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys use affixation to alter call meaning. *PLoS ONE*, 4(11), p. e7808. doi:10.1371/journal.pone.0007808
- Peräkylä, A. (2013.) Conversation analysis in psychotherapy. In Sidnell, J. & Stivers, T. (Eds.), *The handbook of conversation analysis* (pp. 551-574). Chichester: Wiley-Blackwell.
- Piirainen-Marsh, A. (2005). Managing adversarial questioning in broadcast interviews. *Journal of Politeness Research*, 1(2), 193-217.
- Poggi, I. (2013). *Multimodal communication in political speech: Shaping minds and social action : International workshop, political speech 2010, Rome, Italy, November 10-12, 2010, revised selected papers*. Heidelberg: Springer.
- Racine, T. P. (2016). The grounds for animal subjectivity and intersubjectivity. *Animal Sentience*, 1(3). Retrieved October 17, 2018, from <https://animalstudiesrepository.org/animsent/vol1/iss1/3/>.
- Romaniuk, T. (2013). Pursuing answers to questions in broadcast journalism. *Research on Language & Social Interaction*, 46(2), 144-164.
- Sacks, H., Jefferson, G. & Schegloff, E. A. (1995). *Lectures on conversation*. Oxford: Blackwell.
- Schegloff, E. A. & Lerner, G. H. (2009) Beginning to respond: Well-prefaced responses to *wh*-questions. *Research on Language and Social Interaction*, 42(2), 91-115.
- Schegloff, E. A. (1968). Sequencing in conversational openings 1. *American Anthropologist*, 70(6), 1075-1095. doi:10.1525/aa.1968.70.6.02a00030

- Schegloff, E. A. (2007). *Sequence organization in interaction*. Cambridge: Cambridge University Press.
- Sidnell, J. (2010). *Conversation analysis: An introduction*. Chichester: Wiley-Blackwell.
- Slocombe, K. E., Kaller, T., Call, J., & Zuberbühler, K. (2010). Chimpanzees extract social information from agonistic screams. *PloS one*, 5(7), e11473. doi:10.1371/journal.pone.0011473
- Sorjonen, M., Raevaara, L. & Couper-Kuhlen, E. (2017). *Imperative turns at talk: The design of directives in action*. Amsterdam: John Benjamins Publishing Company.
- Sporer, Siegfried L. (2014) Bodily communication and deception. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill and J. Bressemer (Eds.), *Body – Language – Communication: Volume 2 : An international handbook on multimodality in human interaction* (pp. 1913-1921). Berlin/Boston: Walter de Gruyter GmbH.
- Stivers, T. & Sidnell, J. (2005). Introduction: Multimodal interaction. *Semiotica*, 2005(156), 1-20.
- Warren, M. (2006). *Features of naturalness in conversation*. Amsterdam: John Benjamins Publishing Company.
- Zlatev, J. (2008). *The shared mind: Perspectives on intersubjectivity*. Amsterdam: John Benjamins Publishing Company.



## Appendices

### Appendix A

**Transcription Conventions** from Enfield, N. J. & Sidnell, J. (2017) *The concept of action* pp. xix-xxi.

- [      Separate left square brackets, one above the other on two successive lines with utterances by different speakers, indicate a point of overlap onset, whether at the start of an utterance or later.
- ]      Separate right square brackets, one above the other on two successive lines with utterances by different speakers, indicate a point at which two overlapping utterances both end or where one ends while the other continues, or simultaneous moments in overlaps which continue.
- =      Equal signs ordinarily come in pairs, one at the end of a line, and another at the start of the next line or one shortly thereafter. They are used to indicate two things:
  - (1) If the two lines connected by the equal signs are by the same speaker, then there was a single, continuous utterance with no break or pause, which was broken up in order to accommodate the placement of overlapping talk.
  - (2) If the lines connected by two equal signs are by different speakers, then the second followed the first with no discernible silence between them, or was 'latched' to it.
- (1.0) Numbers in parentheses indicate silence, represented in tenths of a second; what is given here in the left margin indicates 1.0 seconds of silence. Silences may be marked either within an utterance or between utterances.
- (.)    A dot in parentheses indicates a 'micropause', hearable, but not readily measurable without instrumentation; ordinarily less than 0.2 of a second
- . ? , Punctuation marks are not used grammatically, but to indicate intonation. The period indicates a falling, or final, intonation contour, not necessarily the end of a sentence. Similarly, a question mark indicates rising intonation, not necessarily a question, and a comma indicates 'continuing' intonation, not necessarily a clause boundary.
- :      Colons are used to indicate the prolongation or stretching of the sound just preceding them. The more colons, the longer the stretching.
- A hyphen after a word or part of a word indicates a cut-off or self-interruption, often done with a glottal or a dental stop.

- word Underlining is used to indicate some form of stress or emphasis, by either increased loudness or higher pitch. The more underlining, the greater the emphasis.
- WOrd Especially loud talk may be indicated by upper case; again, the louder, the more letters in upper case. In extreme cases upper case may be underlined.
- ° The degree sign indicates that the talk following it is markedly quiet or soft.
- °word° When there are two degree signs, the talk between them is markedly softer than the talk around them.
- \_: Combination of underlining and colons are used to indicate intonation contours: If the letter(s) preceding a colon is (are) underlined, then there is an 'inflected' falling intonation contour on the vowel (you can hear the pitch turn downward).
- : If a colon is itself underlined, then there is an inflected rising intonation contour.
- ↓↑ The up and down arrows mark sharper rises or falls in pitch than would be indicated by combinations of colons and underlining, or they mark a whole shift, or resetting, of the pitch register at which the talk is being produced.
- >< The combination of 'more than' and 'less than' symbols indicates that the talk between them is compressed or rushed.
- <> Used in reverse order, they can indicate that a stretch of talk is markedly slowed or drawn out. The 'less than' symbol by itself indicates that the immediately following talk is 'jump-started', i.e. sounds like it starts with a rush.
- hh Hearable aspiration is shown where it occurs in the talk by the letter *h* – the more *h*'s, the more the aspiration. The aspiration may represent breathing, laughter, etc. If it occurs inside the boundaries of a word, it may be enclosed in parentheses in order to set it apart from the sounds of the word.
- .hh If the aspiration is an inhalation, it is shown with a dot before it (usually a raised dot) or raised degree symbol.
- ( ) Empty parentheses indicate that something is being said, but no hearing (or, in some cases, speaker identification) can be achieved.

## Appendix B

**Multimodal transcription conventions** from Mondada (2007) Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies*, 9(2), 195–226.

- \* \* Gestures and descriptions of embodied actions are delimited between
- + + two identical symbols (one per participant)
- Δ Δ and are synchronized with corresponding stretches of talk.
- \*---> The action described continues across subsequent lines
- > until the same symbol is reached.
- >> The action described begins before the excerpt's beginning.
- >> The action described continues after the excerpt's end.
- ..... Action's preparation.
- Action's apex is reached and maintained.
- ,,,,, Action's retraction.
- ric Participant doing the embodied action is identified when (s)he is not the speaker.
- fig The exact moment at which a screen shot has been taken
- # is indicated with a specific symbol showing its position within the turn at talk.